Vol. 63, No. 3 May, 1956

Psychological Review

THEODORE M. NEWCOMB, Editor
UNIVERSITY OF MICHIGAN
Lorraine Bouthilet, Managing Editor

CONTENTS

Operationism and the Concept of Perception.....WENDELL R. GARNER, HAROLD W. HAKE, AND CHARLES W. ERIKSEN 149

Coin Perception Studies and the Concept of Schemata.....HAROLD G. MCCURDY 160

The Nature of the General Reasoning Factor.....J. P. GUILFORD, NORMAN W. KETTNER, AND PAUL R. CHRISTENSEN 169

Phenomenological Personality Theories and the Data Language of Psychology.....RICHARD JESSOR 173

A Formal Theory of Social Power.....JOHN R. P. FRENCH, JR. 181

Drive, Incentive, and Reinforcement.....JOHN P. SEWARD 195

Information-Receiving Behavior of Man.....MASANAO TODA 204

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

Solomon E. Asch	Robert B. MacLeod
Robert R. Blake	David C. McClelland
Stuart W. Cook	George A. Miller
Clyde H. Coombs	Gardner Murphy
Leon Festinger	Oscar Oeser
W. R. Garner	Carl Pfaffmann
James J. Gibson	Carroll C. Pratt
D. O. Hebb	John P. Seward
Harry Helson	David Shakow
E. R. Hilgard	Richard L. Solomon
Carl I. Hovland	Eliot Stellar
E. Lowell Kelly	S. S. Stevens
David Krech	Eric Trist
Robert W. Leeper	Edward L. Walker
Kenneth MacCorquodale	Robert W. White

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STREETS, LANCASTER, PA.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

$8.00 volume $1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 15th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1956 by the American Psychological Association, Inc.

VOL. 63, NO. 3 MAY, 1956

THE PSYCHOLOGICAL REVIEW

OPERATIONISM AND THE CONCEPT OF PERCEPTION [1]

WENDELL R. GARNER, HAROLD W. HAKE, AND CHARLES W. ERIKSEN

The Johns Hopkins University

The attitude of contemporary operationists toward perceptual research has been well characterized recently by Allport. He has described their attitude by stating "that a perception can be regarded as nothing more nor less than a discriminatory response" (1, p. 53). In even simpler terms, the "reaction is the perception," and thus the role of the researcher is simply to determine the conditions under which a discriminatory response is obtained. These conditions then define perception. Unfortunately, we have to agree that many psychologists who consider themselves operationists do in fact accept this position toward perception. However, this position is not necessary from the tenets of operationism. In fact, we believe that this viewpoint is a perversion of the fundamentals of operationism as stated by its originators.

The essence of the above position is that a concept of perception is not distinguishable from the operations on which it is based, and thus that perception is indistinguishable from the responses which indicate its existence and character. This idea springs from a restricted interpretation of Bridgman's writings. For example, Bridgman states that "The concept is synonymous with the corresponding set of operations" (4, p. 5). This widely quoted statement has been used by psychologists to justify their unwillingness to distinguish between perceptions and responses, and to support their position that any set of responses leads to a concept about the properties of the perceptual system. However, to state that a concept is synonymous with a set of operations is not to state that any operation can produce a concept.

Furthermore, psychologists have ignored the fact that Bridgman is talking about a *set* of operations, not a single experimental operation. He later emphasizes this distinction, stating that

> Operational definitions, in spite of their precision, are in application without significance unless the situations to which they are applied are sufficiently developed so that at least two methods are known of getting to the terminus. Definition of a phenomenon by the operations which produced it, taken naked and without further qualification, has an entirely specious precision, because it is a description of a single isolated event (3, p. 248).

Many operationists accept the sterile point of view described by Allport, and consider perception not to have any operationally determinable properties other than discrimination. It seems to us that the above quotations from Bridgman do not require such a narrow point of view. It is true that if the only op-

[1] The preparation of this report was supported in part under Contract N5ori-166, Task Order 1, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-201, under that contract.

eration allowed were the discriminating reaction, then it would be impossible to determine whether the perceptual process had any properties other than discrimination. But it is equally unjustifiable to state that perception is nothing more than the discriminating reaction, since other possibilities have not been excluded. Surely the perceptual process has more richness than simple discrimination.

It seems clear that many operationists who nominally subscribe to the narrow operational position also feel that perception is more than discrimination, since they do in fact ascribe other properties to the perceptual process. It is equally unjustifiable to ascribe such additional properties, when they are ascribed on the basis of single experiments whose designs are inadequate to determine the nature of these additional properties. It is our contention that additional properties of the perceptual system can and must be considered, but that operational experiments of a particular type are necessary to determine the nature of these properties.

PERCEPTION AS A CONCEPT

We conceive of perception as an intervening process between stimuli and responses, as schematically illustrated in Fig. 1. We can directly observe only stimuli and responses and, therefore, perception can be known only as a concept whose properties are induced from

FIG. 1. A schematic of the perceptual problem. The three systems operate in a causal relationship as indicated, although there may be interdependencies. Each system, however, may have independent properties. The properties of the stimulus and response systems can be directly observed; those of the perceptual system must be inferred.

objectively determined relations between stimuli and responses. This statement does not in any sense imply that perception is identical to responses (or to stimuli). Indeed it is the purpose of this paper to indicate the kinds of operations which make it possible to distinguish perception from responses, and to show that these operations are necessary if the concept of perception is to have any use or meaning.

The discriminating reaction. We agree with contemporary operationists that the fundamental and prerequisite operation in any experiment on perception is to demonstrate a discrimination between stimuli on the basis of responses. In other words, it must be demonstrated that there is a contingency relationship between stimuli and responses. If such a relation is demonstrated, we know that a subject can use the same label for the same stimulus within a certain error tolerance.

This operation alone, however, assures us only that we have a system which is operating and which is reliably assigning responses to the various stimuli. This operation provides us with so little information about perception that indeed we cannot distinguish between perceptual and response processes. However, if perception is considered a concept separate from the response system, there are rules for inductively determining the properties of that concept, and they are the same as those for determining the properties of any concept.

Converging operations. The necessary condition which makes possible the determination of particular characteristics of any concept (including the concept of perception) is the use of what have been called converging operations (6). Converging operations may be thought of as any set of two or more experimental operations which allow the selection or elimination of alternative hypotheses or concepts which could

explain an experimental result. They are called converging operations because they are not perfectly correlated and thus can converge on a single concept.

To illustrate, let us assume an experiment similar to one reported by McGinnies (11) in which visual thresholds are determined for words with different emotional content. In our hypothetical experiment, we present four alternative stimulus words, tachistoscopically—fire, save, shit, fuck. The responses to these words are the verbal pronunciation of them. We find that the two vulgar words have higher thresholds than the two nonvulgar words. Such a result has usually been interpreted as indicating that the perceptual system differentially discriminates on the basis of the emotional content of the perceived stimuli. There is, however, at least one other alternative explanation of these results—namely, that the difference in threshold is a result of a characteristic of the response system, which inhibits the verbalizing of some of these words. A converging operation which would discriminate between the two alternative hypotheses would be to present the same stimuli as before, but to pair these stimuli with responses such that vulgar responses are used for nonvulgar stimuli, and vice versa. This experiment in conjunction with that of the first would allow us to decide which of the two hypotheses is correct, or if some combination of them is. It should be noted that the two operations taken together provide the convergence. One experiment does not converge on the other, but rather the two converge on a mutually acceptable result.

Ideally, converging operations would be orthogonal (completely independent), since such operations are the most efficient. In practice, however, it is difficult to obtain truly orthogonal operations, because the world is so organized that all variables cannot be controlled completely independently. This fact does not seriously change the nature of the problem, because a sufficient number of partially converging operations can still provide precise delimitation of alternative concepts.

Nevertheless, sets of operations can be considered more or less efficient in allowing the formation of definitive concepts. One class of operations can be excluded entirely from consideration. This class can be called *parallel operations*. These operations select among alternative hypotheses along the same dimension, and thus cannot converge to a single concept. In our example above, we could do the entire experiment again using different words, but in which two are still vulgar and two nonvulgar. If this experiment produces the same result as the first, we still do not know whether the difference in threshold is due to a characteristic of the perceptual system or of the response system.

There are two special types of parallel operation which deserve mention. One of these is the *repeat operation*, in which the same experiment is repeated at another time. Such an experiment does not allow convergence to a single concept unless the concept involves time as a variable. Another special type of parallel operation is the *transform operation*, in which one variable is simply a transformation of another. Again, in the illustration we have been using, suppose we retain the original set of words as the stimuli, but use four synonyms as responses—burn, keep, crap, screw. These synonyms used as responses can be considered transforms of the original words used as responses, and this experiment will not allow us to determine whether the differential effect exists in the perceptual system or in the response system.

It should be obvious from the above discussion that the value of a set of converging operations depends less on

the nature of the operations themselves than on the alternative hypotheses or properties which are being considered. For example, we mentioned above the use of synonyms as an example of a transform operation. If, however, the synonyms were used as the stimuli, this experiment in conjunction with the first would converge to determine whether the differential thresholds were due to the letter configurations of the original words, rather than to their meanings. Thus these two experiments would be converging for that purpose, but would not be converging for purposes of determining whether the effect was in the perceptual system or in the response system.

Although a minimum of two converging operations may define a concept, in practice it is rare that two are sufficient. In the first place, there are usually more alternative hypotheses than can be delimited with two converging operations. In the second place, converging operations are rarely orthogonal, but are usually only partially converging. If there are enough partially converging operations, we can still arrive at a single, well-defined concept, but more than two such operations will be necessary.

It is quite legitimate to use assumed converging operations in place of operations actually carried out. We know from previous research that many of the possible converging operations would work in a particular way if we actually tried them, and thus there is little point in trying them again. The fact that we use assumptions, however, should not allow us to lose sight of the fact that the validity of our concepts rests entirely on the validity of the operations, whether carried out or assumed. A concept has no meaning other than that derived from the operations on which it is based, and unless these operations are known, the concept cannot be known either.

Operational Distinction Between Response and Perception

An important use of converging operations is to distinguish effects which exist in the response system from those which exist in the perceptual system. Since the primary function of perceptual research is to determine something about the properties of the perceptual system, a major requirement in such research is to use converging operations which will eliminate the possibility that the outcome of an experiment is due to properties of the response system rather than of the perceptual system.

One of the more important properties of the response system which can affect the outcome of a perceptual experiment (and possible interpretations about properties of the perceptual system) is that of *response differentiation*. For example, if only one response is available to a subject, it is clearly impossible to demonstrate anything about perceptual discrimination. This principle is quite obvious when stated in such an extreme form, but it can also operate in other less obvious ways. For example, if the number of response categories is too small to demonstrate the perceptual discrimination capacity of a subject, then the outcome of the experiment will be limited by a property of the response system rather than by a property of the perceptual system. Thus, a converging operation which is required in many experiments is one which demonstrates that discrimination is invariant with respect to the number of response categories and with respect to the discriminability or differentiation of these categories. Such converging operations must either be carried out or be assumed before we can state anything about the limits of perceptual discrimination; if they are assumed, then the validity of the conclusion is limited by the validity of the assumption.

A second major property of the response system which can affect interpretation of experimental results is that of *response availability*. To some extent this property can be considered as a special case of response differentiation. If a subject fails to use a particular response which is defined as appropriate to a particular stimulus, we cannot state that there was a failure of the perceptual system. The failure of response can be due to sheer motor inability. It can also be due to such things as response inhibition (as in the case of vulgar words). That is to say, the response system can be affected by emotional factors just as well as the perceptual system can.

There are other ways in which response availability can operate. For example, it is known that human subjects (and rats too) have a preference for some responses over others. These preferences can affect the apparent nature of the relation between stimuli and responses, and, unless they are taken into account, can lead to misinterpretations about properties of the perceptual system. Subjects also exhibit sequential effects in their responses and, if preferred sequences of responses conflict with the sequences of stimuli presented, misinterpretations can again occur. This factor is commonly taken into consideration in psychophysical experiments, where truly random sequences of stimuli are rarely used. Rather, modified random sequences are used which prevent long runs, since most subjects do not believe that long runs can occur by chance and thus are unwilling to use them.

In summary, then, there are many ways in which properties of the response system can affect the outcome of a perceptual experiment and, unless converging operations are used to delimit an effect specifically to the perceptual system, properties can be incorrectly ascribed to that system.

CONVERGING OPERATIONS AND SOME CURRENT PROBLEMS IN PERCEPTION

To illustrate more specifically the implications of this discussion for research on perception, we shall discuss three types of experiments in these terms, pointing out the kinds of converging operations which are necessary before the results of experiments can reasonably be ascribed to particular properties of the perceptual system.

Subception. The concept of subception has been introduced by Lazarus and McCleary (10) to explain the results of a certain class of experiments on perception. Typically, in these experiments, subjects make two different but simultaneous responses to stimuli from a set presented one at a time. Usually some of the stimuli have a meaning (frequently anxiety-producing) which is different from that of other stimuli in the set. This different meaning may be inherent in the words, or may be induced by conditioning procedures. One response is overt, usually verbal. The other response is often physiological, nonverbalizable, and presumably indicative of emotion. The GSR has been most frequently used. The result which subception purports to explain is that the nonverbalized response shows discrimination when the verbalized response does not. To explain this result it is assumed that perceptual discrimination occurs subconsciously: thus the term subception, which ascribes a property to the perceptual system.

Underlying this definition of subception is the equation of consciousness with verbal report, a distinction that we are reluctant to accept without more explicit specification of the converging operations on which it rests. But even assuming that these converging operations can be specified, the inference of unconscious perceptual discrimination

from the typical subception experiment is not justified. It is not justified because the converging operations are lacking that would permit us to ascribe the results to properties of the perceptual system as opposed to properties of the response system. The fact that the GSR is a nonverbally reportable response, is not sufficient cause for assuming that the perceptual discrimination is also not verbally reportable.

Eriksen (5) has pointed out that the subception result can be looked at basically as the demonstration of a partial correlation between the stimulus and the GSR, with the verbal response held constant. When the subception experiment is conceived this way, many properties of the response system which could produce the subception result become more obvious. For example, Eriksen has noted that such factors as lack of verbal response differentiation, asynchronous response oscillation, and differential response strengths will satisfy the requirements for the partial correlation.

When the problem is looked at this way it becomes obvious that a demonstrated partial correlation as indicated above does not allow us to infer the property of subconscious discrimination to the perceptual system.

Rather, in order to demonstrate that there is perceptual discrimination which cannot be verbally described, it is necessary to show that the first-order correlation of stimuli with nonverbal responses is greater than that between stimuli and verbal responses. But even if a single experiment gave this result, it would still be necessary to use converging operations which demonstrate that the correlation between stimuli and verbal response was not limited by the size or degree of differentiation of the response set. Bricker and Chapanis (2) have shown that more precise discrimination of stimuli can be accomplished verbally if more alternative responses are allowed. If no verbal response system can be found which gives as good a correlation with stimuli as the GSR, then and only then can we conclude that additional perceptual discrimination is operating at a nonverbal level.

Perceptual set. Many experiments have demonstrated that the nature of the relation between stimuli and responses changes when the response system changes. A common interpretation of such experiments is that the set of responses provided the subject serves to produce a perceptually selective set. In many of these experiments, however, the necessary converging operations have not been undertaken to justify delimitation of the effect to the perceptual system.

Suppose an experiment similar to one by Hyman and Hake (9), in which subjects are required to identify the form of stimuli presented tachistoscopically. In one condition subjects are told prior to the stimulus exposure that the form will be one of two particular alternatives. In another condition they are told that the form will be one of four. When the duration of exposure required for identification of the form is shorter with just two stimuli, a possible conclusion is that subjects adopt a more accurate perceptual set with the smaller number of stimuli.

There is, however, an alternative explanation which requires no assumptions about the accuracy of the perceptual process, namely, that the amount of error in the response system decreases with a decrease in number of response categories. One necessary converging operation, then, is to run another experiment in which four stimuli are presented, but only two response categories are allowed. If the accuracy of identification in this case is identical to the two stimulus-two response case, then we must assume that the effect is due entirely to a characteristic

of the response system. To complete the set of converging operations, it would be necessary also to present two stimuli but to allow four response categories. If the responses, however, are direct descriptions of the stimuli, this last operation would be meaningless to carry out.

Actually, there is still another set of converging operations necessary if two responses are used with four stimuli. If the nature of the responses and stimuli are such that the subject can observe one of two aspects of the stimulus, then this operation is parallel to that of using two stimuli and two responses. For example, suppose that the four stimuli were red square, blue square, red circle, blue circle; and the two responses were red, blue. For all practical purposes, there are just two stimuli, not four, and this operation would not allow us to determine where the effect exists. If the responses were abstract, such as letters of the alphabet, and were not assigned to the stimuli by just color or form, the operation would be converging. One system for handling this problem is to use the two responses with all possible pairs of the four stimuli to determine the extent to which meaningful grouping of the stimuli affects discrimination.

Sensory scales. One area of perceptual research where these principles of converging operations are most difficult to apply, and have been applied least, is that of sensory magnitude scales. Psychological scales have been produced for pitch, loudness, weight, brightness, length, and many other perceptual attributes. In most cases it is assumed that the resultant scale tells us about a property of the perceptual system, but we believe that very few of the experiments have included the necessary converging operations to ensure that the final function truly describes a perceptual process.

To illustrate, let us suppose an experiment similar to one reported by Reese *et al.* (12). We provide a large circular, vertical display on which appears a single line radiating from the center. This line is set to various angular positions, and subjects are asked to estimate the angle, using as responses the whole numbers from 0 to 360. When we have obtained many estimates, and averaged them for given physical settings, we plot the estimated angle against the presented angle. At first glance it would appear that we could call such a function a psychological scale of inclination having interval properties. We have no converging operations, however, to indicate that the subjects actually used the numerical response system in such a way as to reflect the perceived magnitudes of angles. People have had considerable experience with such estimates, particularly with clocks, and in the course of this experience they have learned to identify certain perceptions with certain numbers. The fact that later they can do so in an experiment does not alter the fact that they are using the numbers simply as learned identifying responses. Since the subjects have had experience with the 0 to 12 scale but not with the 0 to 360 scale, it might be argued that the new number system would actually reflect properties of the perceptual system. However, as we pointed out above, one operation which is a transform of another does not satisfy the requirements of converging operations.

There is an operation which would appear to satisfy the requirements of convergence for the property under consideration—an interval property of the perceptual system. This operation would be to have the subjects set series of three radiating lines in such a way that two equal angular intervals are provided. In this case a number system is

not used, and thus the chance that the numbers are used simply as identifying labels would seem to be minimized. Subjects are so accustomed to using numbers in this situation, however, that they may very well verbalize the numbers to themselves, and thus once again we have a transform rather than a converging operation.

Actually, in situations where so much learning of objectively assignable numbers to stimulus attributes has occurred, it is doubtful that we can ever sufficiently satisfy the requirements of converging operations to be sure that we are measuring a metric property of the perceptual system, rather than of the objective stimulus system. It would be necessary to get subjects who have had no previous experience in assigning numbers to the continuum under consideration.

Gibson and Bergman (8) have clearly recognized the nature of this problem in their experiments on estimation of distance. Consistent underestimations of distance can be corrected by one or two correction trials (the correction generalizing to all distances used in the experiment), and the authors feel that the correction has simply changed the subjects' conceptual scale of yards. In other words, they are assuming that the change was not in the perceptual system, but rather in the response system. Given a situation like this, we can make statements about the conceptual scale of numbers which the subject has, and its relation to physically measured distance, but we cannot know anything about the metric properties of the perceptual system.

There are other types of psychological dimension that make the problems more obscure because there is no commonly accepted number system which the usual subject has been trained to use. Of these, the one dimension that has received the greatest attention is loudness. Presumably, subjects have had little or no training in using a number scale to describe the loudness of a tone or noise. So when a subject is asked to assign numbers to the loudnesses of various tones, he should be able to respond only on the basis of the perceived magnitudes. This position has been taken by several psychologists, particularly Stevens (13). However, when subjects are required to use a set of responses which constitute a numerical scale, we have no assurance that they in fact use the response continuum in such a way as to reflect the ratio or interval properties of the perceived loudness correctly. The question is not whether they can use the *number* scale correctly, when using it as an abstract scale, but rather whether they use it in such a way as to reflect a metric property of the perceptual system.

Suppose we have done an experiment in which direct numerical estimates of the loudnesses of tones, or of the relative loudnesses of two different tones, are made. From these estimates we can construct a functional relation which we can tentatively call a psychological scale of loudness. Having no assurance that the number scale was used to reflect the metric properties of the perceptual system—the numbers could have been used as simple identifying responses, or perhaps as ordered responses—we need to carry out one or more operations which converge to allow us to delimit a property of the perceptual system. We could ask our subjects to produce tones which sound half, or a quarter, as loud as other tones which the experimenter provides. Such an operation does not converge with the former, however, since the number system whose use is in doubt is the same in both instances.

A considerably better converging operation is one that does not require the use of numbers at all. If a fractionation or direct estimation has been car-

ried out, and if the resultant function is truly a loudness scale with the properties of a ratio scale, then we should be able to determine a set of equal interval points on our scale. Now if we ask our subject to set a series of tones at intensities which provide a series of equal intervals of loudnesses, these equal intervals should check with those obtained from the fractionation experiments. In the one instance in which such a comparison has been made, Garner (7) showed that the results of the two experiments do not check; the two experimental operations did not converge to a single concept about the metric properties of the perceptual system. Garner also showed that if it is assumed only that the fractionation experiments provided equal ratios—not necessarily with the stated magnitude—the two experiments can converge to a single loudness scale. This scale is quite different from that obtained with numerical estimates, however, and its properties and uses are different from those assumed for "scales" obtained from direct estimates.

The major point of this section is not to prove that psychological magnitude scales are impossible to construct but simply to point out that before we can ascribe a property to the perceptual system as refined as that of a ratio scale, certain minimal converging operations are required—operations that in practice have rarely been carried out. Furthermore, it is not sufficient to carry out the same experiment twice, or to carry out a slightly modified experiment which provides only transform operations. The operations must truly be converging for the pertinent assumed property of the perceptual system. Until these operations are carried out, we can have only functional relations, not magnitude scales.[2]

[2] It should be pointed out that there are other scaling techniques for which the ascribed properties of the perceptual system are not at variance with the converging operations carried out. The discriminative scaling techniques, particularly as developed by Thurstone, satisfy these requirements as long as no attempt is made to assume that the resultant scales reflect magnitude of perceptions rather than discriminability of perceptions.

CONCLUSION

Our emphasis on the use of carefully planned operations in order to isolate properties of perceptual behavior should not be taken to mean that operationism properly used tends to minimize the importance of perception. We believe simply that its proper use rigorously defines those properties which validly can be assigned to perception.

This argument holds for all aspects of behavior which can be described as perceptual. For example, the comparison usually made betwen the rich, although unreliable, data provided by introspection and the sterile, although rigorous, data provided by operationism is invalid for two reasons: first, because the basic position taken by many operationists does not indicate the real possibilities inherent in this position; second, and more important, because the data produced by introspection can have no meaning independent of the operations used to produce them.

For example, two aspects of the introspective method have been claimed to have special advantages. First, it is claimed that a much more detailed description of experience can be reported by a subject using unrestricted responses. This problem is not unique with introspection. Operationism does not require that responses be restricted in any way; it only points out that with unrestricted response sets, it is difficult or impossible to separate properties of the response system from those of the perceptual system.

A second claimed advantage of introspection is that subjects can be instructed to confine their attention to

facts of subjective experience, i.e., the stimulus error can be avoided. There is nothing about operationism which excludes this possibility, either. The effects of the subject's set produced by instruction and training are important experimental problems in perception. We insist, though, that if these effects are in the perceptual system, then operations can be devised to demonstrate this fact. These operations must demonstrate that the effect of instructions is not merely to change the kinds of responses that subjects use, nor simply to decrease the correlation between stimuli and responses by causing subjects to respond to factors other than stimuli. That is to say, the methods of introspection are operations, and as such cannot lead to valid concepts that are independent of these operations. Introspection and the data produced by it do not lie outside the scope of operationism as here conceived.

We also feel that there is no need to minimize experimental questions about "awareness." For humans, at least, awareness is undoubtedly one aspect of perception. We insist, however, that the awareness of a subject can be no more than a property of his perceptual system, and that it can be specified only in terms of a set of converging operations.

In summary, we believe that our position is truly one of operationism. We believe that a concept has no meaning beyond that obtained from the operations on which it is based. This statement does not mean that any set of operations can lead to a concept. Nor does it mean that the complexity and usefulness of concepts derived from operations are necessarily limited in any way. In practice they are limited only by the ingenuity of the individual experimenter to devise appropriate converging operations.

SUMMARY

Perception is conceived as a process intervening between stimuli and responses. As such it can be viewed as a concept whose properties may be delimited by converging operations. Converging operations are any set of experimental operations which eliminate alternative hypotheses and which can lead to a concept which is not uniquely identified with any one of the original operations, but is defined by the results of all operations performed. Thus converging operations can lead to concepts of processes which are not directly observable. For example, converging operations can be used to describe properties of the perceptual process which are distinct from those of the response system directly observed.

Illustrations from current experimental problems in perception indicate how some response characteristics may be isolated from perceptual properties, and vice versa. Some of these properties have been ascribed to perception without supporting converging operations by researchers dissatisfied by the sterility of operationism as it is commonly, but mistakenly, conceived.

REFERENCES

1. ALLPORT, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.
2. BRICKER, P. D., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181–188.
3. BRIDGMAN, P. W. Some general principles of operational analysis. *Psychol. Rev.*, 1945, **52**, 246–249.
4. BRIDGMAN, P. W. *The logic of modern physics.* New York: Macmillan, 1927.
5. ERIKSEN, C. W. Subception: fact or artifact? *Psychol. Rev.*, 1956, **63**, 74–80.
6. GARNER, W. R. Context effects and the validity of loudness scales. *J. exp. Psychol.*, 1954, **48**, 218–224.
7. GARNER, W. R. A technique and a scale for loudness measurement. *J. acoust. Soc. Amer.*, 1954, **26**, 73–88.

8. GIBSON, ELEANOR J., & BERGMAN, R. The effect of training on absolute estimation of distance over the ground. *J. exp. Psychol.*, 1954, **48**, 473–482.
9. HYMAN, R., & HAKE, H. W. Form recognition as a function of the number of forms which can be presented for recognition. *USAF, WADC Tech. Rep.*, 1954, No. 54–164.
10. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113–122.
11. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
12. REESE, E. P., REESE, T. W., VOLKMANN, J., & CORBIN, H. H. Psychophysical research: summary report 1946–1952. *USN, Spec. Dev. Cent. Tech. Rep.*, 1953, SDC–131–1–5.
13. STEVENS, S. S. The direct estimation of sensory magnitudes-loudness. *Amer. J. Psychol.*, in press.

(Received October 10, 1955)

Psychological Review
Vol. 63, No. 3, 1956

COIN PERCEPTION STUDIES AND THE CONCEPT OF SCHEMATA

HAROLD G. McCURDY

University of North Carolina

The proposition that perception is influenced by object value and perceiver need has enjoyed an exciting career since it was given prominence by Bruner and Goodman in an oral and, later, a printed report (3) of dramatic differences between rich and poor children in their judgments of coin sizes. Whether that study or subsequent ones can be said to have upheld the proposition may be questionable; but the effect on psychologists is beyond doubt. They were refreshed and stimulated. The ideas, the experimental approach, and the results were new, and quite obviously appealed to current needs.

It was with a little surprise that, perhaps two years ago, while ruffling through the time-blurred pages of the *Psychological Review* of 1906, I ran across an article which seemed to bear some kinship to contemporary studies but which had not, as far as I was aware, been referred to in the recent literature. That accident started a train of thought which, with the aid of a few other bits of historical material, has led to the present paper. Very briefly, what I have to say is this. First, there are certain points of agreement between several coin studies, old and new, but the agreement is not such as to encourage the Bruner-Goodman theory that perception itself is distorted by value and need, at least directly. Second, there may be some advantage in applying the concept of schemata to the coin perception problem, and to other perceptual problems as well. The concept, which goes back to Head and Bartlett, has been very recently discussed in this Journal by Vernon (18).

Old Light on a New Argument

The 1906 paper alluded to above was by Robert MacDougall (11). It tended to show that the symbolic difference between two different denominations of paper money ($1 and $10) is reflected in judgments of their weight. "How many bills together equal the weight of a silver dollar?" was the question MacDougall put to his schoolgirl subjects. To avoid suggestion effects on the main comparison he wished to make, he specified only the one denomination or the other when addressing his query to any particular group. In either case, the correct answer is about 20. The guesses obtained ranged from 5 to a wild 10,000 or so. Not overly concerned about the range, MacDougall thought that the averages and medians favored the view that monetary value has a positive effect on estimates of physical quantity. Some of his data are presented in Table 1.

TABLE 1

Estimates of Number of Bills Needed to Equal Weight of Silver Dollar, MacDougall (10)

School Grade	$1			$10		
	N	Avg	Mdn	*N*	Avg	Mdn
IV	67	215	20	56	125	20
V	155	98	25	116	97	25
VI	78	99	25	66	44	20
VII	40	84	50	42	52	25

MacDougall's inquiry does not stand alone at that period. It was designed, in fact, to follow up a note by J. F. Messenger in *Science* in 1902 (12) which had elicited at least two other brief communications in the same journal, one by Pierce in 1902 (16) and one by Slosson in 1903 (17). All three of these writers offer statistical proof that the weight of paper money is underestimated when compared in memory with coins of gold or silver—a result which may or may not agree with value theories of perception, depending upon how value is defined. But there should certainly be a hearty welcome from that quarter for Messenger's finding that university students and faculty members considered $1 bills lighter than those of higher denomination.

Half a century ago, then, the minds of American psychologists were busy with a question which, in slightly altered form, has concerned us during the last decade. The technique of these early studies may be criticized, and it may be doubted whether the numerical guesses produced by the subjects deserve to be taken literally; but there is a relevance to present-day research on perception of valued objects which should not be ignored simply because the question then was cast in terms of memory judgments unaided by sensory inspection. The problem was the same.

In 1913 G. C. Myers published a monograph on incidental memory (13) which is directly pertinent to recent developments. Among other things he investigated the accuracy with which people estimate the sizes of familiar coins. He presented his numerous subjects with a sheet of cardboard bearing compass-drawn circles ranging in diameter from 9 to 44 millimeters by 1-millimeter steps, and asked them to select from these the six circles corresponding most nearly in size to penny, nickel, dime, quarter, half-dollar, and dollar.[1] In

[1] He considered the circles with diameters of 19, 21, 18, 24, 31, and 38 mm. as fairly representing the sizes of the penny, nickel, dime, quarter, half-dollar, and dollar, respectively, the precise diameters of which he gives on the authority of the U. S. Treasury as 19.05, 21.2, 17.9, 24.25, 30.6, and 38.1 mm.

TABLE 2

ESTIMATES OF DIAMETERS OF COINS (IN MILLIMETERS) BASED ON MEMORY AND VISUAL INSPECTION

Standards		Dime (17.8)	Cent (19.0)	Nickel (21.2)	Quarter (24.1)	Half Dollar (30.5)
Memory						
Myers (N=58–59)		15	17	19.7	25.1	34
Bruner and Goodman	Size=Area?	18.8	19.4	23.2	27.8	35.6
(N=20)	Size=Diam.?	19.9	20.4	25.4	31.9	41.4
Carter and Schooler (N=48)		16.6	18.7	21.1	26.7	34.6
Inspection						
Bruner and Goodman	Size=Area?	20.2	20.4	23.6	28.2	34.6
(N=20)	Size=Diam.?	23.0	22.1	26.3	33.0	39.5
Bruner and Rodrigues (N=30)			19.3	22.7	28	
Carter and Schooler (N=48)		16.4	18.1	21.1	25.6	33.5
Carter and Schooler (N=48)	Cardboard Discs*	16.6	18	21.1	24.4	31.9

* The diameters of the discs were 17.5, 18.7, 21.2, 24.1, 30.6, thus differing slightly from the diameters of the coins.

Table 2, alongside some modern data, I have entered his results for fifth-grade boys and girls, converting his error figures by a simple calculation into average estimated diameters. The agreement of Myers with Carter and Schooler (7) is very close. Though the correspondence with Bruner and Goodman is less close, it is still pretty good, provided that we assume that their "size" equals area rather than diameter—a problem of interpretation to be discussed below.

The features of the data in the first half of Table 2 most deserving our attention are brought out in Fig. 1. The memory estimates from all three studies follow the same slope, deviating from the slope of perfect match (indicated by the heavy line) in such a way that the size range between largest and smallest coins is greater than it would be if the estimates were perfectly accurate. According to Myers and according to Carter and Schooler, the smaller coins are underestimated and the larger overestimated. According to Bruner and Goodman, all sizes are overestimated, but slope and range are not thereby thrown out of agreement with the other two sources of data. In sum, three independent studies of children at about the fifth-grade level agree that in making size estimates of familiar coins from memory the children produce a correct order of coins according to size, and slightly exaggerate the differences between them.

How should these facts be interpreted? Bruner and Rodrigues (5), fighting a kind of rear-guard action against Carter and Schooler in defense of some earlier conclusions, have argued that the relatively greater exaggeration of size in the estimates of coins of larger denomination is further evidence for the hypothesis that symbolic value influences the estimate of physical magnitude. The argument is weak, and it seems to disregard the fact that smaller coins may be underestimated; the dime, for instance, is underestimated, and underestimated to a greater degree than either penny or nickel, according to Myers, and Carter and Schooler, as Fig. 1 and Table 2 show. Should we not say that memory, in stretching out the range of coin sizes, commits a "good" error? If we attend merely to individual coins, we note overestimation here and underestimation there; but if we attend to the whole series, we see that the *order* of sizes is maintained, perhaps guaranteed by the exaggeration of differences. After all, is it not, practically, more important to realize that the dime is smaller than the penny than to know either the absolute sizes or the exact amount of difference between them? One would scarcely deny that symbolic value enters into the construction of such a functional schema. But how? The exaggeration of differences would be more useful than sheer magnification of individual items, and this is what appears to happen.

Perhaps it might be laid down as a general principle that, wherever it is important to have clarity, there will be a tendency to exaggerate differences within any closely graded series, and thus, in a case like the present, lead to a slope steeper than the slope of perfect match. In the case of series of few members and large differences between adjoining members, there might be some lowering of the slope, since this could be afforded without the risk of confusion. But the coin series is a fairly closely graded one, and to keep them separate in memory seems to require some exaggeration of difference.

Before going any farther in the discussion, let me explain a doubtful aspect of Table 2, indicated by the question marks. In trying to put all available information into identical form as millimeters of diameter, I had to treat

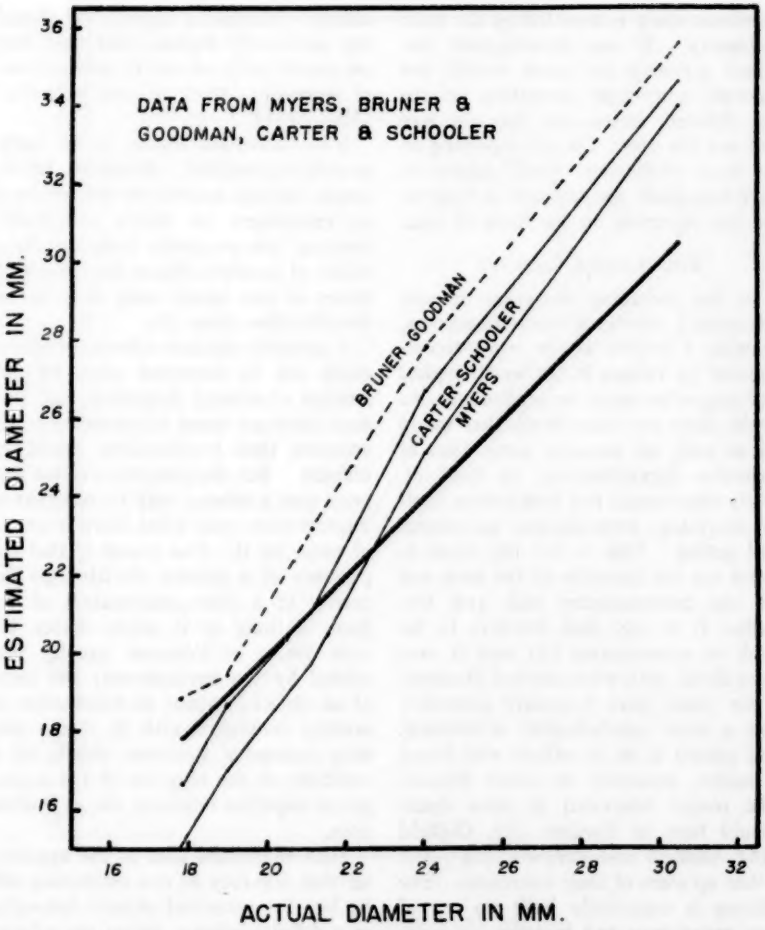


FIG. 1. Accuracy of memory for size of coins as shown by three independent studies.

the percentages of deviation published by Bruner and Goodman according to two different assumptions. Since they are never specific about it, they could mean either area or diameter when they refer to "size." Carter and Schooler apparently take them to mean diameter. The subsequent Bruner and Rodrigues article does not deny that assumption. Nevertheless, it is my guess that "size" in the Bruner and Goodman article means area. My reason is that the diameters derived according to this assumption from their percentage figures harmonize better with the other data in Table 2, including those from Bruner and Rodrigues, than do diameters derived according to the other assumption. Incidentally, the meaning of the percentage figures in the Bruner and

Goodman study is troubled by the same ambiguity. If two investigators obtained precisely the same results but reported percentage deviations on the two different bases, one the one way and one the other, the one reporting on the basis of diameter would appear to have deviations roughly half as large as the one reporting on the basis of area.

THE SCHEMA CONCEPT

In the preceding discussion of size estimates I once used the term schema, in what I believe is the sense recommended by Vernon in her recent topical and suggestive paper on perception. No doubt there are other terms that would do as well, all meaning some kind of cognitive disposition not in itself directly experienced but manifesting itself in imagining, remembering, perceiving, and acting. This is not the place to ferret out the ancestors of the term and all the contemporary kith and kin. Suffice it to say that Bartlett in his book on remembering (1) took it over from Head, with whom he had discussed it for years, gave it greater generality and a more psychological orientation, and passed it on to others who found it useful, especially in Great Britain. The reader interested in more detail should turn to Bartlett (1), Oldfield (14), Oldfield and Zangwill (15), and follow up some of their references. The schema is organically built up out of past experiences and reactions, it functions as an orienting and stabilizing factor in new contacts with the environment, and, although it is constantly subject to change both because of organismic dynamics and environmental variations, it may come to govern behavior quite rigidly in exceptional cases. As Vernon says, "The essential point to note is that the use of a concept such as that of the schema reflects the fact that our percepts, thoughts, and behavior are on the whole consistent and orderly—reasonably appropriate though not necessarily logical—and that they are consistently related to previous acts of perceiving, thinking, and behaving" (18, p. 181).

Now, schemata appear to be inadequately appreciated. Brunswik, for example, betrays surprise at the results of an experiment by Bolles and Bailey showing how extremely little the direct vision of familiar objects improves estimates of size based upon mere verbal identification alone (6).

A perfectly accurate schema, of course, could not be improved upon by any amount of sensory inspection, and perhaps there are many schemata so nearly accurate that improvement would be difficult. But everyday observation suggests that a schema may be resistant to improvement even when there is plenty of room for it. One reason is that the presence of a schema obviates the necessity of a close examination of objects, so long as it works within the wide limits of tolerance usually permitted by the environment; the name of an object is about as informative as sensory commerce with it, if the sensory commerce functions chiefly as a reminder to the observer of the appropriate cognitive reference file, as it often does.

Such reflections lead to the hypothesis that accuracy of size estimation will be less for perceived objects belonging to a definite schema, unless the schema has been very carefully developed, than for perceived objects lacking such a schema. Figure 2 supports this hypothesis with data drawn from Carter and Schooler. Inspection of the coins does improve the estimates based upon the schema operating in memory, but the estimates for the cardboard discs, which presumably are less schematized than the coins, are more accurate still. The evidence is admittedly slender, but, such as it is, it is not contradicted by

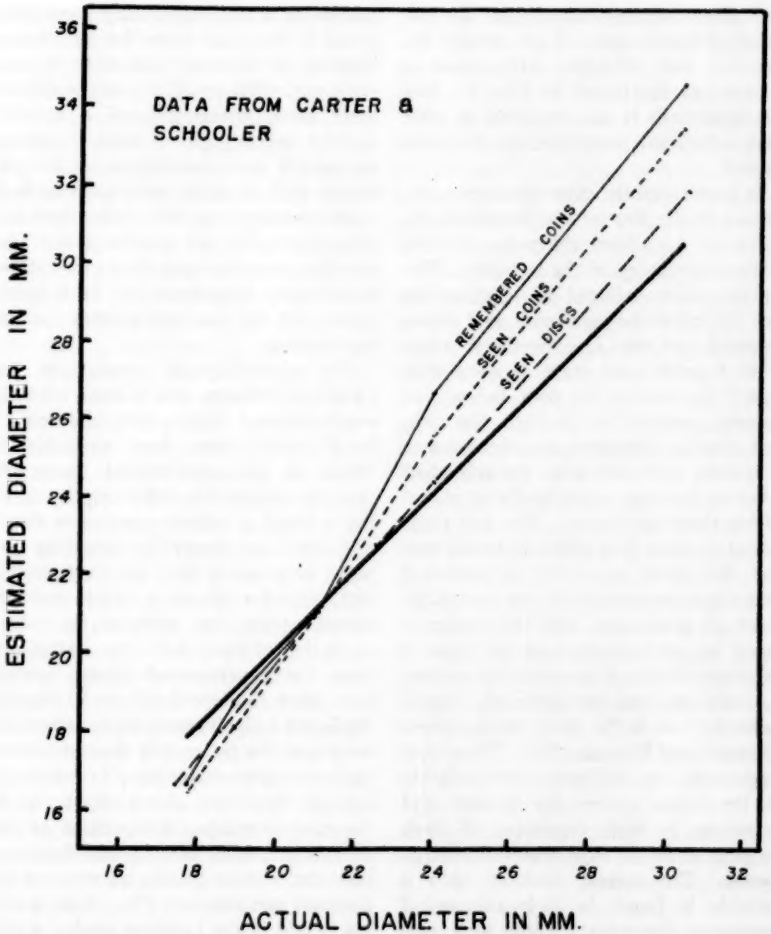


FIG. 2. Comparative accuracy of memory and immediate perception.

the data from Bruner and Goodman; if anything it is only too well supported. One expects sensory inspection to result in some improvement where the memory estimate is inaccurate; but, contrary to this expectation, just the opposite happened in the Bruner and Goodman study, as nearly everybody knows. On the other hand, the estimation of cardboard discs, both in this study and in the later one by Bruner and Rodrigues, was very accurate indeed.

EMOTIONAL EFFECTS?

I have argued that under ordinary circumstances one would expect estimates of size to be progressively better as one passes from an estimate based upon a memory schema, through one representing the interaction of schema

and direct sensory inspection, to one based primarily upon direct sensory inspection and relatively little upon a schema, as illustrated in Fig. 2. But the hypothesis is not prepared to cope with emotional perturbations, however caused.

It is not exactly clear what the conditions of the Bruner and Goodman experiment were from the point of view of the psychology of the subjects. The children were selected as representing two distinct economic classes, and it was assumed that the class membership implied a predictable state of need with regard to money, the poor being supposedly greedier for it than the rich, and tending, therefore, to overestimate coin sizes, especially when the coins had taken on burning reality by being placed within their very grasp. But this theoretical position is modified in the Bruner and Rodrigues paper; it is confessed that class membership is an unreliable index of need state, and the reader is urged to put more trust in the type of approach "where a money-value system is built up and extinguished experimentally," as in the study by Lambert, Solomon, and Watson (10). There was, apparently, no deliberate manipulation of the value system by Bruner and Goodman in their treatment of their subjects or in the experimental arrangements. The merest hint of such a variable is found in their manner of presenting the coins. These were presented, verbally and materially, in ascending and descending *order of value* rather than randomly or in order of size. Whether this procedure accounts for the slight tendency to dislocate the dime from its proper relative position in the size series is an open question. As to producing a general overestimation of size for all the coins, however, it would scarcely suffice. Were the children emotionally aroused in any way by the experimenters when the coins were presented, as might conceivably have happened if there had been, for instance, a promise of monetary reward? Apparently not. The conditions of the experiment were, presumably, of a neutral, strictly psychophysical kind, implying no reward and arousing no excitement except such as might have been derived spontaneously from the need states appropriate to the rich and the poor. The startling results, now that the class-membership hypothesis has been abandoned, are left hanging without specific explanation.

The possibility of emotion in the Lambert, Solomon, and Watson study is much clearer. Very young children (3 to 5 years) were used as subjects. Those in the experimental group repeatedly obtained a poker chip by turning a crank a certain number of times and then got candy by inserting the poker chip into a slot; the control subjects likewise turned a crank and received candy, but without the intermediation of the poker chip. After ten days the experimental group, having been often reinforced (10 or 50 times), displayed a significantly higher size estimation of the poker chip than in pretest trials or after extinction; interestingly enough, there was also a slight rise in the control group. A repetition of the experiment, with certain modifications, produced similar results, as reported by Lambert and Lambert (9). Now, nothing is said in the Lambert studies about the emotions of these little children. Nevertheless, it may be fair to guess that the process of reinforcement generated some degree of pleasant excitement, whereas the extinction experience may have been somewhat sobering. The former emotional condition might have been more conducive to an expansive sort of inaccuracy, especially in the presence of the stimulating poker chip, than the latter. At any rate, it is not unreasonable to look for an explanation

of the expanded estimates in the general pleasant emotion induced in both sets of children by the interesting and rewarding game, and especially induced in the experimental group by the fact that the poker chip which they had to judge was an intrinsic part of the game, suggesting by its presence still more candy or other exciting outcome. Perhaps this is what is meant by Lambert and Lambert when commenting on a variation of the experiment in these words: "Still another way of discussing these data is that there is an effect at least analogous to stimulus intensity involved. That is, the white token is a stronger stimulus, at least in its visceral prededications, and it results in an amplification of the judgment response" (9, p. 509). One should not expect children of 3 to 5 years of age to be very critical in making the sensorimotor adjustments required in these experiments for indicating their judgments, when they are under the stress of pleasant excitement. That the error happens to come out in the direction of a larger estimate rather than a smaller may be an expression of an affective law, though at the moment I can think of no better evidence for such a law than the familiar rough observations concerning expansion of the body and bodily gestures in sthenic states, fluctuations of handwriting size with changes of mood, the general overplus of activity in excited children, and so on.

SUMMARY AND CONCLUDING REMARKS

The following principal conclusions emerge from the present discussion: (*a*) The memory schema for coin sizes, while stretching out the total range and slightly exaggerating the differences between coins, preserves the correct order; (*b*) there is no evidence that this order is significantly disturbed by values and needs; (*c*) a memory schema, such as that for coin size, tends to resist correction by sensory contact with objects, contact serving as little more than a fleeting reminder, like a name; (*d*) departures from accurate size estimation which have been taken to favor the hypothesis that increase of value induces phenomenal magnification might better be interpreted as the result of the interference of expansive emotional states with meticulous adjustments.

The above conclusions refer to certain specific studies and to results stated in the form of averages. There is no intent to deny that values and needs may influence perception in *some* way, though presumably more through schemata than directly. The experimental evidence is weak, however, and very fragmentary.

One final point may be made. If a strong case is ever to be developed for need and value effects on perception, it may be supposed that it will have to be founded on the careful analysis of individual data. Perceivers are always individuals, and schemata, too, are individual matters, even though there may be percepts and schemata which are widely shared in a society or species. Perhaps it is not scientifically desirable to restrict ourselves so narrowly to average data, to neat tabulations of measurements and tests of significance, while neglecting to describe fully the general behavior of our subjects, and neglecting, in fact, to study the pattern of data for each individual. Klein, Schlesinger, and Meister (8) have done us a service, I believe, in pointing out the overwhelming importance of the individual differences in their own value-perception study, and have properly emphasized that mass statistics may completely hide the value determinants which may be operating quite powerfully in individuals taken singly.

REFERENCES

1. BARTLETT, F. C. *Remembering.* Cambridge: Univer. Press, 1932.

2. BRUNER, J. S. One kind of perception: a reply to Professor Luchins. *Psychol. Rev.*, 1951, **58**, 306–312.
3. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1947, **42**, 33–44.
4. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Pers.*, 1949, **18**, 14–31.
5. BRUNER, J. S., & RODRIGUES, J. S. Some determinants of apparent size. *J. abnorm. soc. Psychol.*, 1953, **48**, 17–24.
6. BRUNSWIK, E. In defense of probabilistic functionalism: a reply. *Psychol. Rev.*, 1955, **62**, 236–242.
7. CARTER, L. F., & SCHOOLER, K. Value, need, and other factors in perception. *Psychol. Rev.*, 1949, **56**, 200–207.
8. KLEIN, G. S., SCHLESINGER, H. J., & MEISTER, D. E. The effect of personal values on perception: an experimental critique. *Psychol. Rev.*, 1951, **58**, 96–112.
9. LAMBERT, W. W., & LAMBERT, E. C. Some indirect effects of reward on children's size estimation. *J. abnorm. soc. Psychol.*, 1953, **48**, 507–510.
10. LAMBERT, W. W., SOLOMON, R. L., & WATSON, P. D. Reinforcement and extinction as factors in size estimation. *J. exp. Psychol.*, 1949, **39**, 637–641.
11. MACDOUGALL, R. On secondary bias in objective judgment. *Psychol. Rev.*, 1906, **13**, 97–120.
12. MESSENGER, J. F. How many one-dollar bills will equal in weight a five-dollar gold piece? *Science*, 1902, **15**, 672–673.
13. MYERS, G. C. A study in incidental memory. *Arch. Psychol.*, 1913, **4**, No. 26, 1–108.
14. OLDFIELD, R. C. Memory mechanisms and the theory of schemata. *Brit. J. Psychol.*, 1954, **45**, 14–23.
15. OLDFIELD, R. C., & ZANGWILL, O. L. Head's concept of the schema and its application in contemporary British psychology. Part I. Head's concept of the schema. Part II. Critical analysis of Head's theory. Part III. Bartlett's theory of memory. Part IV. Wolters' theory of thinking. *Brit. J. Psychol.*, 1942, **32**, 267–286; **33**, 58–64, 113–129; 1943, **33**, 143–149.
16. PIERCE, A. H. Guesses on the relative weights of bills and coins. *Science*, 1902, **16**, 745.
17. SLOSSON, E. E. Guesses on the relative weights of bills and coins. *Science*, 1903, **17**, 189.
18. VERNON, M. D. The functions of schemata in perceiving. *Psychol. Rev.*, 1955, **62**, 180–192.

(Received June 6, 1955)

Psychological Review
Vol. 63, No. 3, 1956

THE NATURE OF THE GENERAL REASONING FACTOR[1]

J. P. GUILFORD, NORMAN W. KETTNER, AND PAUL R. CHRISTENSEN

University of Southern California

In the report on primary mental abilities found in his first large factor analysis, Thurstone listed a factor whose unique property he thought to be an ability to reason under restrictive conditions (11). Its largest loading was in a test of arithmetic reasoning. The same factor was not subsequently identified as such in his later analyses.

In the AAF Aviation Psychology Program research during World War II, a factor that was found repeatedly as a consistent feature of arithmetic-reasoning tests was called *general reasoning* (7). Because of the unusual variety of tests found related to the factor, no clear definition of it was achieved. It was concluded that the ability is a kind of "general-purpose ability" or "trouble-shooting ability," which seems to come into play in a difficult test or one that contains some very difficult items.

A hypothesis concerning this factor that would be appealing to some is that it is synonymous with general intelligence. Finding it in a large variety of tests would be one support for this idea. The fact that an arithmetic-reasoning test is a popular component of many intelligence tests would be another. Against the hypothesis, however, is a very stubborn fact. Analysis shows that many types of intelligence tests are not related to the factor; for some types, correlations are near zero.

No one has seriously maintained that the factor is the only ability to reason, for a number of reasoning factors have been found. Our own results show that there are as many as five that can be called reasoning factors.

At the beginning of the project studies of aptitudes at the University of Southern California, we favored the hypothesis that *general reasoning* has something significant to do with problem solving. Although not all problems are cast in arithmetical form, and there are indeed a great variety of problems, we tended to follow traditional thinking to the effect that underlying the solving of problems in general there is a common psychological core—that we can speak of problem solving in a generic sense. We have to admit that we do not know much about the nature of this phenomenon, but *general reasoning* might be one of the common elements.

Our first analysis of reasoning abilities began with five alternative hypotheses regarding the nature of *general reasoning* (2). The factor might be:

1. A general ability to manipulate symbols. This would make the ability perhaps as broad as thinking itself, thus going beyond problem solving.

2. A general ability to solve problems. This narrows the conception to a more reasonable level.

3. An ability to define problems. This concentrates attention on the first steps of problem solving, and stops short of the actual solving activities.

4. An ability to test hypotheses. Hypothesis testing occurs during the defining of a problem and also at later stages.

5. An ability to organize a sequence of related steps. An arithmetic-reasoning problem that is at all complex requires a sequence of numerical operations to solve it. We know that the actual numerical operations depend upon

[1] Based upon a paper presented at the annual convention of the American Psychological Association in San Francisco, September 1–7, 1955.

the factor of *numerical facility*. An arithmetic-reasoning test is commonly loaded on both this and the *general-reasoning* factor.

Our first two factor analyses (2, 5) eliminated the first hypothesis, in agreement with previous analyses. Not all tests involving symbols are loaded on it. Furthermore, we found a new ability to manipulate symbols on a restricted scale. That is, it was confined characteristically to tests whose items approach equation form.

A recent analysis of planning abilities (4) has shown a factor that could be called "ordering ability," which in a way takes care of Hypothesis 5, for the factor is separate and distinct from *general reasoning*.

Results have more and more pointed to the elimination of Hypothesis 2, that the factor is a general problem-solving ability. There are too many tests involving problems in which the factor is absent. This is especially true where the test material is composed of figures rather than words and/or numbers. We have come to the view that a common, unique, *psychological* core for all problem solving does not exist (3). Problems are simply too varied, and each type seems to call upon its own pattern of abilities—perceptual abilities as well as thinking abilities.

As we have gone from one analysis to another, our preferred hypothesis concerning the factor has undergone changes, without reaching a final resolution. The interpretation that was favored after each successive analysis can be briefly stated:

6. Effectiveness of trial-and-error manipulation of symbols in the solving of problems of certain kinds (2). This gives much attention to the fact that problems that are so difficult as to involve the examinee in somewhat random attempts are the best material for measurement of the factor.

7. Ability to consider a number of conditions in defining or understanding a problem (5).

8. Ability to size up a test (or problem) situation to see what is demanded by it (12).

9. Ability to structure a problem (8).

10. Ability to find and maintain a problem structure when considerations are complex (4).

Examination of these suggestions shows that the original Hypothesis 3, that the factor is an ability to define problems, is very persistent. We also see that two new ideas have entered the picture. One emphasizes a possible trial-and-error aspect and the other emphasizes the complexity of the task. The latter idea suggests a span concept. Individuals may differ with respect to the level of complexity they can handle in thinking. The complexity may be calibrated ideally in terms of the number of elements, relations, or variables that the problem entails.

A very recent analysis (6), which was concentrated on this factor, attempted to reach some decision among the three alternatives; that the factor is an ability (*a*) to understand or define a problem, (*b*) to think effectively by trial and error, and (*c*) to handle a complex set of elements, relations, or variables.

Tests designed to measure ability to understand problems concentrated on the first stages of solving arithmetic-reasoning problems. Given a question, what facts are or are not essential to finding an answer? Given a problem, what arithmetical operations are essential? Tests emphasizing trial and error presented problems in which trying one thing after another would be a natural mode of attacking the task. Tests emphasizing complexity either required the handling of many elements in every item or increased the number progressively throughout the test. It will have to be admitted that experimental sepa-

ration of these features between sets of tests was not completely achieved.

The results throw considerable doubt on the trial-and-error hypothesis but do not eliminate it from all future consideration. The three tests especially designed to emphasize trial and error were not loaded on the factor.

One of the four tests designed to emphasize complexity had a relatively high loading on the factor, in fact the highest for any test in this analysis. This test is known as Ship Destination. In each item the problem is to tell how many miles a ship is from a port on a simple map. New conditions that modify the effective distance are added as the test progresses—wind direction, current direction, initial heading, and strength of wind and of current. In a sense, the test belongs in the category with arithmetic reasoning. Structuring the problem would still be important.

Two of the three tests designed to emphasize definition of problems had substantial loadings on the factor. The third had shown some relationship to the factor in two earlier analyses. It had been revised in the meantime, however, with an effort to slant the test toward some other factor. Perhaps this effort had succeeded too well.

Thus, we are forced back upon the hypothesis having to do with understanding problems. Just what kind of understanding is important and what kinds of problems, in addition to those of the arithmetic-reasoning type, are still to be determined.

During the course of our investigations, other hypotheses have been suggested from other sources. In his monograph on aptitude factors, French calls the factor "deduction" (1). There are one or two other reasoning factors that have greater claim to this name and that have come out in analyses, distinct from *general reasoning*. This hypothesis is thus easily eliminated.

More recently Lucas and French (10) have suggested that the factor is the ability to do quantitative or mathematical reasoning. Under this hypothesis, it would be difficult to explain why the factor is consistently loaded in tests that present problems of a nonmathematical and nonquantitative type (such as reading comprehension) and in which it is difficult to see how numbers and mathematical thinking would help. We also have evidence from our project research that there is sometimes little or no relation of *general reasoning* to criteria of achievement in college mathematics (9). We must therefore also reject this hypothesis.

In conclusion, we may say that it has been much easier to decide what *general reasoning* is not than to say what it is. We have given reasons for believing that it is not general intelligence, though it is undoubtedly one of the important intellectual abilities; it is not a broad ability to manipulate symbols, but there seems to be a narrow ability of this sort; it is not a generic ability to solve problems, though it undoubtedly plays a role in solving problems of certain types; it is not an ability to order a sequence of steps, but there is an ordering ability, which may be of some importance in planning; it is not an ability to use trial and error in problem solving; and it is not a general idea span, though the span concept may be found to apply in some limited way yet to be determined.

By elimination and by consistent indications of a positive nature, the best we can say is that *general reasoning* has something to do with comprehending or structuring problems of certain kinds in in preparation for solving them. The range of problems that are pertinent is still to be determined. Whether this ability extends beyond problem solving is also to be determined. It may be a

general ability to formulate complex conceptions of many kinds.

REFERENCES

1. FRENCH, J. W. The description of aptitude and achievement tests in terms of rotated factors. *Psychometr. Monogr.*, 1951, No. 5.
2. GREEN, R. F., GUILFORD, J. P., CHRISTENSEN, P. R., & COMREY, A. L. A factor-analytic study of reasoning abilities. *Psychometrika*, 1953, **18**, 135–160.
3. GUILFORD, J. P. Factors in problem solving. *ARTC Instructors J.*, 1954, **4**, 197–204.
4. GUILFORD, J. P., BERGER, R. M., & CHRISTENSEN, P. R. A factor-analytic study of planning, II. Administration of tests and analysis of results. *Psychol. Lab. Rep.*, Los Angeles: Univer. of Southern Calif., 1955, No. 12.
5. GUILFORD, J. P., GREEN, R. F., CHRISTENSEN, P. R., HERTZKA, A. F., & KETTNER, N. W. A factor-analytic study of Navy reasoning tests with the Air Force Classification Battery. *Educ. psychol. Measmt*, 1954, **14**, 301–325.
6. GUILFORD, J. P., KETTNER, N. W., & CHRISTENSEN, P. R. A factor-analytic investigation of the factor called general reasoning. *Psychol. Lab. Rep.* Los Angeles: Univer. of Southern Calif., 1955, No. 14.
7. GUILFORD, J. P., & LACEY, J. I. (Eds.) *Printed classification tests.* Washington, D. C.: Government Printing Office, 1947. (*AAF Aviat. Psychol. Program Res. Rep.*, No. 5.)
8. HERTZKA, A. F., GUILFORD, J. P., CHRISTENSEN, P. R., & BERGER, R. M. A factor-analytic study of evaluative abilities. *Educ. psychol. Measmt*, 1954, **14**, 581–597.
9. HILLS, J. R. The relationship between certain factor-analyzed abilities and success in college mathematics. *Psychol. Lab. Rep.*, Los Angeles: Univer. of Southern Calif., 1955, No. 15.
10. LUCAS, C. M., & FRENCH, J. W. *The factorial composition of the relative movement test.* Princeton: Educational Testing Service, 1953.
11. THURSTONE, L. L. Primary mental abilities. *Psychometr. Monogr.*, 1938, No. 1.
12. WILSON, R. C., GUILFORD, J. P., CHRISTENSEN, P. R. & LEWIS, D. J. A factor-analytic study of creative-thinking abilities. *Psychometrika*, 1954, **19**, 297–311.

(Received October 2, 1955)

Psychological Review
Vol. 63, No. 3, 1956

PHENOMENOLOGICAL PERSONALITY THEORIES AND THE DATA LANGUAGE OF PSYCHOLOGY

RICHARD JESSOR

University of Colorado

It is clear that an increasing number of theories of personality find it useful, if not necessary, to employ some construct referring to the *psychological* situation —e.g., life space (10), phenomenal field (13, 15), or meaningful environment (14). Such theories do not attempt to relate behavior to the physical or geographic environment. Instead, they attempt to make predictions by reference to the environment as it is perceptually, cognitively, or functionally responded to by an organism or class of organisms. In view of the generally recognized importance of such theoretical efforts, on the one hand, and in view of certain questions that have been raised about the methodological status of these efforts, on the other, it would seem opportune to examine the issues involved in some detail.

The most frequently cited challenges to phenomenological[1] orientations in personality theory have questioned the degree to which such orientations may be considered to be truly physicalistic, and the degree to which they may be considered to be predictive rather than postdictive systems. These issues are, of course, not unrelated to each other. Lewin's system, perhaps because of the explicitness with which it is presented, has served most often as the prototype of phenomenological theories when these issues have been discussed. Brunswik (4), for example, describes Lewin's life space as postperceptual and prebehavioral, and doubts whether Lewin's predictions can, in a strict sense of the word, be tested. Spence (16) has stressed the failure of such field theories as Lewin's to provide us with laws that will enable control and manipulation of the behavior-determining psychological field. His analysis has gone further to emphasize the distinction between the types of laws that are achieved by phenomenological theories versus those attained by systems which do refer behavior to the physico-geographic environment. The former laws are called R-R laws, and are considered to relate one set of responses to another set of responses obtained earlier from the subject or the experimenter. The latter laws are referred to as S-R laws, and are considered to relate responses to independently measured physical and social environmental variables.

The aims of this paper are threefold: to support the position that phenomenological theories of personality are compatible with the general scientific requirements of physicalism and predictiveness, to demonstrate that a fundamental methodological distinction between these theories and nonphenomenological theories is difficult to maintain, and, finally, to suggest that a crucial issue intrinsic to a discussion of the problems mentioned above—and one rarely raised in that connection—is the nature of the psychological data language traditionally favored by S-R and by R-R theories.

[1] The term *phenomenological* and its variants are used in this paper to refer to theories which employ some construct referring to the psychological situation, e.g., life space, behavioral environment, etc. The present usage divests the term of certain historical connotations, such as holism or introspectionism, which have accrued to it.

THE PROBLEM OF THE PSYCHOLOGICAL ENVIRONMENT

The essential characteristic of all constructs which refer to a psychological environment is that their nature and properties are response-inferred. It is unfortunate that some usages of such response inferences have been of the sort as to couple them closely in time to the behavior that is being predicted. This has given the impression of a very limited kind of predictiveness, if not, in actuality, a kind of postdiction. It has also raised the related question of the *independence* of the two sets of responses. Since the problem of prediction and the problem of independence of variables are both closely bound up with the problem of the definition of the stimulus, the latter will be reviewed at some length. In this way a clearer picture of the methodological status of the psychological-environment construct should emerge. It may be noted, parenthetically, that Lewin's theory, as an example of phenomenological theories, *is* considered by Spence (16) to achieve laws between independent variables; the emphasis of his remarks is on the point that the variables are all response variables.

The term *stimulus* in psychology implies, by definition, a relatedness to responses rather than independence of them. It is impossible to isolate, point to, or describe aspects of the environment as stimuli except insofar as they have some determinable relationship to the responses of some organism. To conceive, therefore, of stimulus variables as independent of response variables in general appears to be unwarranted. This point has been raised in connection with other problems by both Brown (2) and Bakan (1). Bakan has noted that variables are simply sets of categories, and *"categories are the result of someone's delineation, abstraction, and identification"* (1, p. 47). In this sense there are no stimulus variables which are not response-inferred or defined, *including the variables of physics and sociology.* The organism which is the reference point for the process of definition in physics is, of course, the experimenter. It follows from these considerations that all stimuli are response-defined at some time by some organism. The issue becomes that of *whose* response definitions, under what conditions.

There are two other ways in which the meaning of the term *independence* may be considered.

1. A distinction may be attempted between the responses of the subjects and those of the experimenter, that is to say, the stimulus definition by the experimenter may be considered independent of the responses of his subjects. Davis (8) has presented this position in its extreme in his recent call for a more physical psychology. Davis states: "For a 'stimulus' (external event) to qualify under the proposed canon, it would have to be something which an experimenter could ascertain without there being any organism for it to work on" (8, p. 10). This statement is worth examining closely. In view of our previous discussion, it may be noted that Davis concedes, at least, that the stimulus is defined by the response of the experimenter, i.e., "something which the experimenter could ascertain." More appropriate to the present point, however, is the fact that the word "stimulus" is placed in quotes by Davis. Obviously, this evidences recognition that external events may be considered stimuli *only* when some relationship is demonstrable between the presence or occurrence of the event and some behavior or response of an organism.

If this view were not held, in actuality, by psychologists, the result would

be the investigation and manipulation of an infinite array of physically discriminable variables (discriminated by the experimenter), some of which may have effects on organism—e.g., a loud noise —and some of which may not—e.g., a sound of a frequency too high to be directly perceived by human subjects. Even the traditional S-R theorists manipulate variables which are defined by the responses of their rats rather than by themselves alone. The empirical law of effect is a concept which illustrates how classes of noxious stimuli and incentives are defined by the approach or avoidance behavior of populations of rats. To summarize this point, it is possible for experimenters to discriminate or identify variables independently of the responses of their subjects, but these variables may be considered *as stimuli* only when they are functionally related to, i.e., defined by, the responses of their subjects.

2. The second sense in which the independence of the stimulus may be considered is in terms of its independence from a given or particular response. Stated otherwise, the issue is whether the stimulus may be identified *before* the particular response occurs. This is a crucial issue, since it involves the entire process of prediction and control. Most of the phenomenological systems have often been content to identify the stimulus only *after* the given response has occurred and, more seriously, have not adequately provided or described techniques or procedures for enabling the experimenter to *predict* behavior by knowledge of the nature of the stimulus in advance of the occurrence of the particular response. Spence (16) is correct in his criticism of such phenomenological approaches for not telling us what to do to an individual in order to manipulate, change, or control his behavior. This criticism, however, applies to phenomenologcial systems such as those of Lewin, Snygg and Combs, and Rogers, *not* because they utilize response-inferred constructs but rather *because they have traditionally and uniformly rejected a historical approach to the problem of prediction.* It is only an ahistorical phenomenology which has limited predictiveness, or is essentially postdictive. A historically oriented phenomenology (e.g., 14) enables the experimenter to predict behavior by providing a basis for advance knowledge of the stimulus.

Let us examine this latter point in greater detail. The phenomenologists have maintained that behavior is a function of the psychological situation (10), the stimulus functions (9), or the meanings of stimuli (14), etc. The problem for them is to know, in advance of the given response, how the subject will constitute the stimulus situation, i.e., what meaning it will have for him. Since the meaning of a stimulus, or the precise way in which a variable functions as a stimulus, is inferred from the responses of the subject, it follows that, for the prediction of a given response in a particular situation, *the stimuli must be inferred from previous responses of the subject,* or similar subjects, in similar or systematically related situations.

This is precisely the way in which S-R behavior theorists proceed to define their stimuli in advance of the behavior of their rats. The way in which it is predicted that a pellet of food in a T maze will be an incentive stimulus for a hungry rat, and will result in approach and eating behavior, is simply by having observed that this rat, or other similarly hungry rats, *previously* approached and ate the food pellet.

For a similar approach to the problems of prediction in personality, it is possible to mention several well-known procedures which enable specification of

the stimulus situation independent of and prior to a given response.

1. *Verbal report, by either the subject or the experimenter.* Spence has pointed out that ". . . the phenomenological approach has its advantages, particularly in the complex field of social behavior of the human adult. It is obviously much easier to gain some notion as to the relevant variables determining such complex behavior by asking the individual to verbalize than it is to employ the procedure of trying to hypothesize them from knowledge of past history" (16, p. 57). Unique to human beings is their ability to verbalize their perceptions of situations or the meanings that they have for them. This information, in spite of its well-known limitations, is exceedingly valuable to the personality theorist interested in predicting the subject's subsequent behavior in those situations. And as Brunswik (5) has pointed out, it is possible to utilize verbal reports without falling back upon introspectionism.

Psychological tests may be considered as a specific example of this general procedure. Test data are essentially self reports or verbalized perceptions of stimulus situations. They are highly useful in attaining knowledge of the meaning or perception of future or subsequent real-life situations by the subject, to the extent that the test situation is similar to or systematically coordinated to the criterion or predicted situation.

2. *Observation of the responses of the general culture group to which a given subject belongs.* It is possible to define a stimulus situation for a given subject on the basis of previous responses by the culture group to which the subject belongs. Obviously, the accuracy or the applicability of such a culture-group definition for a given subject depends upon his relatedness to that group, or the similarity of his past experiences to those of the group members. Snygg and Combs (15), for example, point out that people who share common roles in a common culture develop common characters in their phenomenal fields and consequently in their behavior. In short, similar social learning will result in similar perceptions of environments.

It will be seen that the broader the reference group, the less will be the likelihood that its definition of the situation will be useful or accurate for a *given* individual. If one were interested in the meaning of a classroom quiz situation for a particular college sophomore who is also a fraternity member, the responses of the reference group of college sophomore fraternity members might be more applicable than those of college students as a whole, which in turn might be more accurate (that is, predictive) than those of the middle class as a whole, etc.

3. *The clinical method.* This leads to the third procedure, which assists us in coordinating idiographic and nomothetic approaches. In the preceding paragraph the use of culture-group definitions of situations was advocated. Since such definitions will apply only *more or less* to given subjects, they are useful for predictions about differences between *groups* of individuals, but will be unsatisfactory where specific individual prediction of behavior is desired. In the latter case it becomes necessary to reduce the reference points to the unique past experiences of the particular subject. By intensive study of the subject's previous responses in various situations, the personality theorist is able to establish the meaning or potential perception of a situation for a specific subject, and hence to predict subsequent behavior. This procedure, of course, is the clinical method, and the present description of it emphasizes its continuity with nomothetic procedures.

More concrete elaboration of the implication of these three general procedures for defining situations will be made shortly.

The preceding discussion has stressed the point that all stimuli are response-defined at some time by either the subject or the experimenter. In view of this homogeneity, it seems unwarranted to establish a fundamental logical or methodological distinction between S-R and R-R theories. Viewed in this light, the use of constructs referring to a psychological environment would seem to require no special justification. The discussion to this point has also delineated at least the potential compatibility of phenomenological theories with the necessary demand for predictive adequacy.

THE NEED FOR A PSYCHOLOGICAL DATA LANGUAGE

Granting the argument to this point does not at all lead to the conclusion that S-R and R-R theories do not differ in some significant way. The question to be asked is, Wherein is there a difference between these types of theories? The answer seems to lie in the data language that is favored to describe the situation or environment in which behavior takes place.

There is common agreement among various theorists that psychology thus far has neglected the development of an adequate descriptive terminology for the environment. While personality theorists have been interested for a long time in categorizing the behavioral or internal states of human subjects, they have paid far less heed to developing categories for describing different kinds of situations. Such attempts as have been made are suggestive, but remain limited in usage to particular theoretical orientations, and no widely accepted or universal descriptive terminology for the environment exists at present. It seems to the present author that one difference between the S-R theories and the phenomenological or R-R theories is that the former have heavily employed the data language of physics (despite the reliance on such terms as *cue* and *goal*), whereas the latter have tended toward a more psychological data language. The reliance by the former on physical terms seems to be one of the factors which has contributed importantly to the *appearance* of independent stimulus definition that was discussed above.

Undeniably effective use has been made by S-R theory of the language of physics. Such a data language has the tremendous advantage of being one in which, for historical reasons, maximum reliability or interobserver agreement can be attained, and which involves concepts for which accurate measures have been developed. Nevertheless, reliability alone is insufficient qualification for any language system. What is required is that it be adequate for the purposes and problems of a given science or level of description. Since the real world is neutral with respect to language systems, there is no a priori reason, beyond degree of reliability, for the use of the data language of physics by another science such as psychology. The remainder of this paper will attempt to point out possible reasons why the language of physics has proved unsatisfactory for personality theory, or for psychologists dealing with complex human social behavior, and proposes the development of a psychological data language.

It may forestall misunderstanding to state at this point that our quarrel is not with methodological physicalism, i.e., the insistence on objectivity via denotative reduction or observational reliability. Rather, the critique is aimed at the incorporation into psychology of the *language* of physics. The latter procedure is part of what Brunswik has designated as thematic physicalism,

". . . the uncritical emulation of the . . . aims and problem content of physics by other disciplines" (5, p. 14).

Methodological physicalism in no way implies that the various branches of science must use the same language of description, more specifically the language of physics. Carnap (6) indicates recognition of the necessity for terms other than physical ones in his distinction between physics and biology. He notes, with respect to the latter science, that "The terms which are used in this field in addition to logico-mathematical and physical terms may be called . . . *biological terms*" (6, p. 412). Such terms, as well as psychological terms to describe situations, must, of course, be reducible to observable thing-predicates, i.e., to terms which designate properties which can be determined by direct observation. It is in this sense that Carnap discusses the unity of the language of science, while recognizing the absence of unity of scientific laws. It is congruent with such a position to call for laws which are clearly psychological, i.e., statements of relations between *psychological* terms.

The problem of a data language other than a purely physical one is much more acute for the phenomenological theories dealing with humans than for the S-R theories which have dealt largely with lower organisms. (*a*) The situations in which human social behavior takes place are far more complex than the typical rat-learning situation. For the latter it has been adequate thus far to employ the language of physics, since the manipulated stimulus variables are often simple physical dimensions such as amount of illumination, intensity of electric shock, etc., and since highly restricted and controlled situations are characteristically employed in such research. We do not imply that S-R approaches have utilized only the language of physics; we are maintaining that a large number of terms from that discipline are employed, and that when extension to complex human social situations is undertaken, such terms will prove inadequate, and more appropriate language will have to be invented. How to describe the situation where a college student is asking a coed for a date is an example; description in terms of the amount of illumination, atmospheric pressure, etc., will contribute little to a predictive analysis of such a complex situation. The Hawthorne studies (12) provide an excellent example of how a physical description of the stimulus situation in terms of changes in illumination was inadequate to account for the direction of changes in worker productivity. (*b*) Relatively greater variability of the human response repertoire would seem to be another reason for failure of the language of physics in personality theory. A wide variety of potential responses implies a multiplicity of potential definitions of the same situation by different subjects. (*c*) The presence in humans of language and memory, which allow for the symbolization of variables not physically present in the situation though instigated by the present situation, would seem to be a further complicating factor.

It is proposed that one of the immediately pressing tasks for psychology to undertake is the development of an adequate psychological data language to describe the environment. This is a language whose terms have reference to the nature (behavior) of some organism or class of organisms, since the existence and nature of stimuli depend upon the characteristics, needs, habits, expectancies, intelligence, etc., of organisms. It represents a different abstraction from and organization of the basic referential level, or level of observable thing-predicates, than does the data language of physics. A simple ex-

ample is the term *goal,* which presupposes a directed or striving organism. The same situation, as described physically, would be described differently in psychological terms for different organisms or classes of organisms. Some tentative beginnings in this direction are referred to below.

Rotter (14) has proposed that situations be described by their cultural meanings in terms of the characteristic reinforcements or goals which are likely to occur in those situations. Since no situation always provides just one kind of a goal, situations may be characterized as mixed, or may be described in terms of the dominant or usual or most frequent goals likely to occur. Situations may also be described as ambiguous or unfamiliar where the expectancies for the occurrence of any particular goal are low. Thus, for prediction at the level of personality, situations may be described as love and affection situations, the college quiz may be characterized as an academic recognition situation, etc. Rotter goes on to state:

> In this sense people, too, may be thought of as situations; it makes good sense to speak of authority figure situations, heterosexual situations, and so forth when these terms imply that a particular kind of reinforcement is likely to occur (14, p. 202).

Murray (11) has presented a set of terms belonging to the category of Press, which describes the environment or stimulus situation according to the kind of effect—facilitating or obstructing—that it is exerting or could exert upon the organism. Lewin's (10) terms, such as valence, barrier, etc., are also of this sort.

Chein (7), too, has suggested a geobehavioral language, i.e., a set of concepts to describe the real world looked at from a point of view that is concerned with understanding behavior.

With these examples in mind, the question arises as to the reliability or intercommunicability of such descriptions. But this would seem to be the *only* question which legitimately may be raised, despite any feelings of discomfort over these psychological terms. That intersubjectivity of such descriptions can be achieved may be illustrated from a host of psychological research. Mention here may be made of a series of researches (14) carried out in relation to Rotter's social learning theory, in which characterizations of experimental situations in terms of the potential goals they involve for college students have led to predictions of behavior which have consistently received empirical support.

The development of a psychological language of description might well be enhanced by following the suggestion of McClelland (3, p. 145) that a phenomenological census be carried out to discover what things and attributes in the environment people look for and attend to in guiding their behavior. Fruitful concepts may also be derived from sociology and anthropology, which disciplines have concentrated heavily upon descriptions, with respect to human behavior, of social or cultural situations. Such descriptive terms will, of course, vary in their relevance for different culture groups and especially for different individuals.

It is the increased and explicit utilization of a psychological data language which enables a *rapprochement* between S-R and R-R or phenomenological theories of personality. Although it may be maintained that all psychological laws are fundamentally R-R laws, nevertheless it does make better sense to speak of S-R laws as the goal of psychology when the definition of S is made in *psychological* terms, and the issue of independence of S is restricted to independence from a particular response in a given situation.

REFERENCES

1. BAKAN, D. Learning and the scientific enterprise. *Psychol. Rev.*, 1953, **60**, 45–49.
2. BROWN, D. R. Stimulus-similarity and the anchoring of subjective scales. *Amer. J. Psychol.*, 1953, **66**, 199–214.
3. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake, & G. V. Ramsey (Eds.), *Perception, an approach to personality*. New York: Ronald, 1951. Pp. 121–147.
4. BRUNSWIK, E. Organismic achievement and environmental probability. *Psychol. Rev.*, 1943, **50**, 255–272.
5. BRUNSWIK, E. The conceptual framework of psychology. Chicago: Univer. of Chicago Press, 1952. (*Int. Encycl. unified Sci.*, v. 1, no. 10.)
6. CARNAP, R. Logical foundations of the unity of science. In H. Feigl, & W. Sellars (Eds.), *Readings in philosophical analysis*. New York: Appleton-Century-Crofts, 1949. Pp. 408–423.
7. CHEIN, I. The environment as a determinant of behavior. *J. soc. Psychol.*, 1954, **39**, 115–127.
8. DAVIS, R. C. Physical psychology. *Psychol. Rev.*, 1953, **60**, 7–14.
9. KANTOR, J. R. *Principles of psychology*, Vol. 1. New York: Knopf, 1924.
10. LEWIN, K. *Dynamic theory of personality*. New York: McGraw-Hill, 1935.
11. MURRAY, H. A. *Explorations in personality*. New York: Oxford Univer. Press, 1938.
12. ROETHLISBERGER, F. J. *Management and morale*. Cambridge: Harvard Univer. Press, 1944.
13. ROGERS, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.
14. ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.
15. SNYGG, D., & COMBS, A. W. *Individual behavior*. New York: Harper, 1949.
16. SPENCE, K. W. The nature of theory construction in contemporary psychology. *Psychol. Rev.*, 1944, **51**, 47–68.

(Received July 20, 1955)

Psychological Review
Vol. 63, No. 3, 1956

A FORMAL THEORY OF SOCIAL POWER

JOHN R. P. FRENCH, JR.

Research Center for Group Dynamics, University of Michigan

This formal theory is a small part of the later stages of a program of empirical research on social influence.[1] It tries to integrate previous findings into a logically consistent theory from which one can derive testable hypotheses to guide future research.[2] The more specific purpose of the theory is to explore the extent to which the influence process in groups can be explained in terms of patterns of interpersonal relations.

In discussing the effects of the majority on conformity by the individual deviate, Asch states, "The effects obtained are not the result of a summation of influences proceeding from each member of the group; it is necessary to conceive the results as being relationally determined" (15, p. 186). Both Heider (18) and Newcomb (29) have treated patterns of opinion and of interpersonal relations as a single system of relations, though they have discussed only two-person groups. The present theory reduces the process of influence in *N*-person groups to a summation of interpersonal influences which takes into account three complex patterns of relations: (*a*) the power relations among members of the group, (*b*) the communication networks or patterns of interaction in the group, and (*c*) the relations among opinions within the group. Thus propositions which have been conceptualized at the group level (e.g., that the strength of group standards increases with increasing cohesiveness of the group) are deduced from concepts at the interpersonal level.

The deductive power and the internal consistency of a mathematical model stem from a set of explicit definitions and postulates stated with enough precision so that one can apply the rules of logic. But the construction of theory by coordinating mathematical definitions and postulates to psychological constructs and assumptions leads to a dilemma: the very precision which gives power to the theory also tends to oversimplify it. For reasons of mathematical convenience one tends to make simple assumptions which so restrict the theory that it may seem unrealistic compared to the complexity observed in social behavior. Game theory, for example, describes certain aspects of how "the rational economic man" ought to behave, but actual economic behavior often departs widely from this simple ideal (20).

The present theory deals with this dilemma partly by utilizing a kind of mathematics, the theory of directed graphs, which does not require the making of precise quantitative assumptions about empirical variables.[3] In addition, the basic concepts and postulates

[1] The work reported in this paper was financed in part by a grant from the Rockefeller Foundation and by a contract with the Group Psychology Branch of the Office of Naval Research.

[2] Similar current attempts to construct mathematical theories of social influence include unpublished papers by Ardie Lubin, by Harold Guetzkow and Herbert Simon, and by Solomon Goldberg.

[3] The theory of directed graphs, which is an extension of graph theory (16), has been studied by Frank Harary and Robert Norman with a view toward utilization by social scientists. A publication of this work is planned for the near future (17). The author is indebted to these mathematicians for specific help in proving the theorems of this theory as well as for their work on the theory of digraphs upon which it is based.

of this theory were chosen to conform to the results of experiments on social influence. Frequently, however, our present knowledge was not adequate for making these choices in precise detail. At these points we attempted to choose postulates which would be essentially correct in their main outlines even though some details would have to be changed as new empirical knowledge accumulates. It is not surprising, therefore, that many of the theorems are quite similar to previous findings about influence on opinions and attitudes, even though no research has been done specifically to test this theory. Nevertheless we have intentionally oversimplified the process of social influence by omitting many important determinants and by making very restrictive assumptions about others. It seemed wise to start by examining the implications of a small number of postulates before proceeding to more complex theories.

THE MODEL

Following the theory of quasi-stationary equilibria of Lewin (23), changes in opinion, attitude, or judgment are conceptualized in terms of forces operating along a unidimensional continuum (5, 12). Social influences are coordinated to force fields induced by person *A* on person *B;* and the strength of these forces is assumed to vary with the power of *A* over *B*. The potential force field corresponding to this power relation will be actualized only if *A* communicates to *B* or interacts with him. When *A* expresses his opinion or argues for it in a way that influences *B*, then the force field operating on *B* has a central position corresponding to *A*'s position along the continuum of opinion. All the forces operating on *B* are directed toward this central position, so *B* will tend to change his opinion in a direction which brings him closer to *A*. Similarly, other members, *C, D, E,* etc., who communicate to *B* may set up force fields on him with central positions corresponding to their own opinions. The actual changes in *B*'s opinion will be in accordance with the resultant force from all these induced forces plus a force corresponding to his own resistance.

In order to derive the exact amount of influence that each member will have on the opinion of every other, let us assume that we are dealing with a unidimensional continuum of opinion which can be measured with a ratio scale. We might think, for example, of the classic experiment on social norms by Sherif (32), where the members of the group were asked to state their opinions about how many inches the light moved as they viewed the autokinetic effect. We shall denote the members of the group by *A, B, C,* . . . and their initial opinions by *a, b, c,* . . . respectively, where *a* is the distance of *A*'s opinion from the zero point on the scale. The abscissa of Fig. 1 shows such a scale of opinion

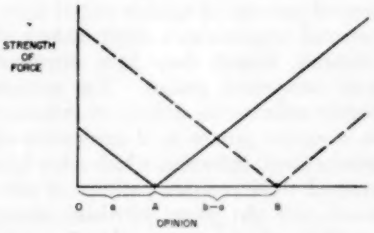


FIG. 1. The force fields influencing opinion.

together with the initial opinions of *A* and *B*. The ordinate indicates the strength of the forces. The gradient of forces around *A* represents the forces he can induce on *B* to agree with his opinion, while the gradient of forces around *B* represent his tendency to resist changing his opinion. Where these two gradients intersect, a distance of $\frac{1}{2}(a + b)$ from the origin, there is an equilibrium point where the two forces are equal in strength and opposite in direction. At

all points to the right of this equilibrium the forces induced by *A* are stronger than *B*'s resistance, so *B* will move toward the point of equilibrium. Conversely, at all points to the left of the equilibrium, *B*'s resistance forces are stronger than *A*'s inductions, so *B* will still move toward the point of equilibrium. Similar calculations of the resultant force and consequent changes of opinion can be made for *A* and for groups with any number of members by placing all members on the same scale and by assuming that the gradient of forces around each member represents both forces he can induce on others and forces he can set up as resistance against others.

The process of influence in a group takes place gradually over a period of time. As one member changes his position and begins to influence others toward his new position, the force fields corresponding to his influence will also shift their central positions. It will be convenient, therefore, to divide the influence process into a sequence of units defined in terms of opinion change rather than in terms of physical time. A *unit* is defined as the time required for all members who are being influenced to shift their opinions to the point of equilibrium of all the forces operating at the beginning of that unit. At the end of the unit, after this shift has taken place, we assume that the members now start to argue for their new opinions. It should be noted that this definition implies that all members respond at the same rate to the forces impinging on them. One possible operational definition of a unit might be a single trial in an experiment such as Sherif's.

This conception of influence as a process over time implies a distinction between direct and indirect influence. In a typical organization the president usually influences indirectly a person at the bottom of the chain of command through orders which are handed down through several subordinates. *Direct* influence is exerted on another person by direct communication which is not channelled through a third person. *Indirect influence* is exerted on another through the medium of one or more other persons. Therefore the direct influence of *A* on *B* always occurs during the same unit, whereas indirect influence requires two or more units. For example, *A* influences *B* directly during the first unit, and *B* influences *C* toward his new opinion during the second unit. Thus *A* has indirectly influenced *C* by transmitting his opinion via *B*. In this model, the power structure and the communication channels of the group are translated into a process of influence over time. In the first unit any member, *A*, influences only those recipients of his communication over whom he has direct power; in the second unit *A*'s influence is also transmitted to all those over whom these intermediaries have power; in the third unit *A*'s influence is transmitted to those who are three steps removed from him in the power structure, etc.

THE POSTULATES

Three main postulates are involved in this model. The first is concerned with interpersonal power. The definition of *power* used in this postulate is the same as that given by Cartwright (6): the power of *A* over *B* (with respect to a given opinion) is equal to the maximum force which *A* can induce on *B* minus the maximum resisting force which *B* can mobilize in the opposite direction.

The *basis* of interpersonal power is defined as the more or less enduring relationship between *A* and *B* which gives rise to the power. French and Raven (in an unpublished manuscript) have discussed five bases: *attraction power* based on *B*'s liking for *A*, *expert power*

based on B's perception that A has superior knowledge and information, *reward power* based on A's ability to mediate rewards for B, *coercive power* based on A's ability to mediate punishments for B, and *legitimate power* based on B's belief that A has a right to prescribe his behavior or opinions. Any basis of power can vary in strength: there may be variations in how much B likes A, in how much B respects A's expertness, etc. Postulate 1 is general enough to refer to all bases of social power.

Postulate 1. For any given discrepancy of opinion between A and B, the strength of the resultant force which an inducer A can exert on an inducee B, in the direction of agreeing with A's opinion, is proportional to the strength of the bases of power of A over B.

Attraction as a basis for interpersonal influence has been demonstrated in experiments by Back (2) and by French and Snyder (13), and in field studies by Lippitt, Polansky, and Rosen (25). Expertness as a basis for interpersonal power has been demonstrated in the latter two studies as well as in many others (19, 27, 28). In unpublished experiments French and Raven and French, Levinger, and Morrison have demonstrated that legitimacy and the ability to punish are bases for social power. Heider (18) and Newcomb (29) state their theories in terms of "positive relations," a more general conception which combines several types of power. In most real groups the power relations probably do combine several of the bases discussed here and others too. Postulate 1 refers to all of these bases combined.

Resistance, as a part of the social power discussed in Postulate 1, has not been treated separately nor in detail in this model. In a further development it might be coordinated to such factors as "certainty of own opinion" (13), or as Kelman (21) and Mausner (26) call it, "prior reinforcement," and to various personality characteristics such as rigidity and authoritarianism (8).

Postulate 2. The strength of the force which an inducer A exerts on an inducee B, in the direction of agreeing with A's opinion, is proportional to the size of the discrepancy between their opinions.

This postulate combines two effects which have been demonstrated in previous research. (*a*) More influence is attempted toward the member who is more discrepant (10, 11, 30). These studies also show, however, that this effect holds only under conditions where the inducee is not rejected. Too great a deviation leads to changes in the attraction power structure of the group and hence to changes in the effects implied by Postulate 1. (*b*) If the amount of influence attempted is held constant, the amount of change in the inducee increases with increasing size of discrepancy. For this latter relation, French and Gyr (12) report correlations of .77, .62, .65, and .83 in different experimental groups. Goldberg (14) also reports a strong tendency for the amount of change to increase with increasing discrepancy, with the inducee moving 30 per cent of the way toward the inducer for discrepancies of all sizes. In a subsequent unpublished theoretical paper,[4] Goldberg also assumes that change in opinion is a direct function of discrepancy until the inducee rejects the credibility (expert power) of the inducer, after which it becomes an inverse function of discrepancy. Again the data support Postulate 2 within the range where the expert power structure of the group is not changed.

Postulate 2 is represented in Fig. 1 by the two increasing gradients of forces

[4] S. C. Goldberg, Some cognitive aspects of social influence: a hypothesis. (Mimeographed.)

around *A*'s opinion and around *B*'s opinion. The two gradients are assumed to be linear, though the evidence cited above would suggest that they are curvilinear. We have made the more convenient assumption because it appears to be true as a first approximation and because it seems to be possible to revise the postulate later, if subsequent empirical data do show curvilinearity, with only minor quantitative changes in the theorems.

Postulate 3. In one unit, each person who is being influenced will change his opinion until he reaches the equilibrium point where the resultant force (of the forces induced by other members at the beginning of the unit and the resisting force corresponding to his own resistance to change) is equal to zero.

Postulate 3 is an application of a basic assumption of Lewin (22) that locomotion or restructuring will take place in the direction of the resultant force whenever that force is greater than zero. Though consistent with a great many empirical studies, this assumption is close to a conceptual definition which cannot be directly tested.

THEOREMS

For lack of space, no attempt will be made to state all the theorems which have been proven nor to give the formal proofs of those presented. Instead we will select some representative theorems and indicate informally the nature of the derivations. In making empirical predictions from these theorems, this theory, like any other, must always assume "other things being equal," including all extrasystem influences and the many factors within the group which are not part of the theory.

The Effects of the Power Structure of the Group

This section presents some theorems concerning the effects of the power structure of the group on the influence process and its outcome. These theorems illustrate how the present theory explains a well known proposition about groups in terms of concepts about interpersonal relations.

This proposition—that the strength of group standards increases with increasing cohesiveness of the group—has been substantiated in several studies (2, 4, 9, 30, 31). A group standard has been defined conceptually as group-induced pressures toward uniformity of behavior or belief, and it may be measured by the degree of conformity of members produced by these pressures. *Cohesiveness* has been defined conceptually as the resultant forces on members to belong to the group, but it has been operationalized in many of these experiments as the attraction of members for one another (7, 24). Festinger, Schachter, and Back (9) have shown that the hypothesized relation is stronger when cohesiveness is operationalized in a way which takes account of the pattern of the sociometric structure instead of a simple summation of choices. But each sociometric choice measuring the attraction of member *B* toward member *A* is, according to Postulate 1, a basis for *A*'s power over *B*. Thus the sociogram of a group can be transformed into the attraction power structure of the group by simply reversing the direction of each arrow. The attraction power structure of the group is a special type of power structure, and hence it is treated in these theorems about power structure and trends toward uniformity of opinion within the group.

The power structure of a group may be represented conceptually in terms of the mathematical theory of directed graphs, called "digraphs." A digraph is a finite set of points $A, B, C, \ldots$ and a subset of the directed lines $\overrightarrow{AB}, \overrightarrow{BA}, \overrightarrow{AC}, \overrightarrow{CA}, \overrightarrow{BC}, \overrightarrow{CB}, \ldots$ between distinct points. In representing

power structures as digraphs, we shall coordinate points to members and directed lines to power relations between members. In this coordination we shall make only relatively crude distinctions in differences of power: if "*A* has power over *B*," there is a directed line $\overrightarrow{AB}$ in the digraph representing the power structure of the group; if "*A* does not have power over *B*," there is no such line.

Various properties of digraphs may be used to characterize power structures of groups. We shall be concerned here primarily with the "degree of connectedness" of power structures. In order to discuss this property we need two definitions: *complete digraph* and *directed path*. A digraph is complete if there exists a directed line from each point to every other point. A power structure would be complete, then, if each member had power over each other member. If we assume that when *A* chooses *B* sociometrically *B* has power over *A*, then it follows that when every member of a group chooses every other member, the digraph representing the power structure of the group will be complete (e.g., No. 5 in Fig. 2). A *directed path* is a collection of distinct points *A*, *B*, *C*, . . . , together with the lines $\overrightarrow{AB}$, $\overrightarrow{BC}$, . . . If in the power structure of a group there is a directed path from *A* to *C*, it follows that *A* can exert influence on *C* even though *A* may not have direct power over *C* (there must be a sequence of directed lines originating at *A* and going to *C* even though there is no line $\overrightarrow{AC}$).

In their work on digraphs Harary and Norman (17) have defined four degrees of connectedness. Their definitions are as follows: (*a*) A digraph is *strongly connected* (or *strong*) if for every pair of distinct points, *A* and *B*, there exists a directed path from *A* to *B and* a directed path from *B* to *A*. It follows that every complete digraph is strong, but not every strong digraph is complete. (*b*) A digraph is *unilaterally connected* (or *unilateral*) if for every pair of points, *A* and *B*, there is a directed path from *A* to *B or* from *B* to *A*. (*c*) A digraph is *weakly connected* (or *weak*) if it is impossible to separate the points of the digraph into two classes such that no line of the digraph has one end point in one class and the other end point in the other class. Thus, for every possible separation of all of the points of a weak digraph into two disjoint, nonempty classes, there must be at least one line having one end point in one class and the other end point in the other class. (*d*) A digraph is *disconnected* if it is not weak. Thus a disconnected digraph may be separated into two (or more) disjoint classes of points such that no line goes from one class to the other. From these definitions it is clear that all strong digraphs are unilaterally and weakly connected and that all unilateral digraphs are weakly connected. It is also clear that all weak digraphs are *not* strongly connected. For this reason it is useful to define a digraph as *strictly unilateral* if it is unilateral but not strong, and to define a digraph as *strictly weak* if it is weak but not unilateral. In our discussion here, when we speak of unilateral or weak digraphs we shall mean "strictly unilateral" and "strictly weak."

In groups where each member communicates to all others over whom he has direct power during every unit of the influence process, the amount of uniformity achieved and the speed of achieving it tend to vary with the degrees of connectedness of the power structure, except that no differences were proved for weak vs. disconnected digraphs.

The effect is illustrated in Fig. 2, and generalized later in the first four theo-

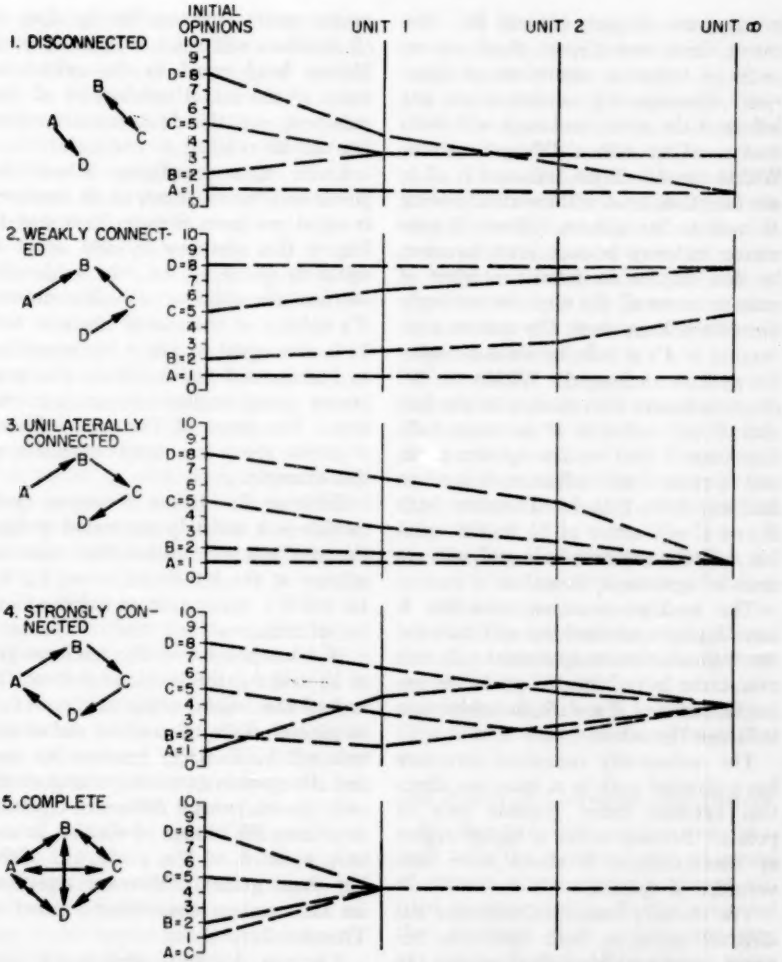


FIG. 2. The effects of connectedness on opinion changes in the group.

rems. In Fig. 2 five different types of structures (complete, strong, unilateral, weak, and disconnected) are illustrated by digraphs of four-person groups. To the right of each structure are curves showing some of the theoretically predicted changes of opinion. The ordinate gives the scale of opinion and, at the left, the initial opinion of members *A*, *B*, *C*, and *D*. The line labeled "Unit 1" gives the distribution of opinion after the first unit; the line labeled "Unit 2" gives the distribution of opinion after the second unit; and the line labeled "Unit ∞" gives the equilibrium of opinions reached in an infinite number of units.

The disconnected structure is com-

posed of two cliques, *AD* and *BC*. Between these two cliques there are no paths of influence, regardless of direction. Consequently neither clique can influence the other, and each will eventually end up with a different opinion. Within the *AD* clique, influence is all in one direction, so *A* will eventually swing *D* over to his opinion. Since *D* only moves half-way in each unit, however, he will require an infinite number of units to move all the way. Accordingly the dotted lines show *D*'s opinion converging to *A*'s at infinity while *A*'s opinion remains unchanged. Within the *BC* clique, influence is mutual, so in the first unit *B* will influence *C* to move halfway from 5 to 2 on the opinion scale, and likewise *C* will influence *B* to move half-way from 2 to 5. Therefore both *B* and *C* will arrive at $3\frac{1}{2}$ on the opinion scale in the first unit and will remain in agreement thereafter.

The weakly connected structure is more highly connected but still does not result in unanimous agreement. In this case there is no directed path, for example, between *A* and *D*, so neither can influence the other.

The unilaterally connected structure has a directed path in at least one direction between every possible pair of points. Because it has a higher degree of connectedness, it shows more convergence of opinion.

The strongly connected structure has directed paths in both directions between every possible pair of points. In this example the strongly connected digraph is a cycle, yielding a final common opinion which reflects more equal influence of all members.

The completely connected structure has direct, one-step paths in both directions between the members of every possible pair. It converges in only one unit to a final common opinion.

Theorem 1. For all possible patterns of initial opinion, in a completely connected power structure the opinions of all members will reach a common equilibrium level equal to the arithmetic mean of the initial opinions of all the members, and this final common opinion will be reached in one unit.

Under these conditions where the power and the resistance of all members is equal, we have already illustrated in Fig. 1 that the new opinion of *A* is equal to $\frac{1}{2}(a + b)$, i.e., the arithmetic mean of the opinions of both members. *B*'s opinion at the end of the first unit b_1 is also equal to $\frac{1}{2}(a + b)$, according to Postulates 2 and 3. Thus this two-person group reaches agreement in one step. The proof of Theorem 1 for an *N*-person group is a simple extension of this example.

Theorem 2. In an *N*-person cycle (which is a strongly connected group) the members will reach a final common opinion at the arithmetic mean, $(1/N)(a + b + c + \ldots)$, in an infinite number of units.

If *A* has power over *B*, then $b_1 = \frac{1}{2}(a + b)$ and b_2, the opinion of *B* at the end of the second unit, $= \frac{1}{2}(a_1 + b_1)$. In general, *B*'s opinion at the end of any unit will be half-way between his own and *A*'s opinion at the beginning of the unit; so the general difference equation describing *B*'s change of opinion in any unit, n, is: $b_n = \frac{1}{2}(a_{n-1} + b_{n-1})$. Solving these general difference equations for all members constitutes a proof of Theorem 2.

Theorem 3. In a unilaterally connected group the opinions of all members will converge to a final common opinion in an infinite number of steps.

It is an obvious theorem of digraph theory that no strictly unilateral digraph can have more than one point of input zero, i.e., with no directed lines leading to it (because then these two or more points could not have a directed path between them—which violates the definition of a unilateral di-

graph). It follows that, during every unit, at least one of the two members at the extremes of the range of opinion will be subject to the power of another and will move toward the center, thus restricting the range of opinion still further. Eventually, therefore, all members will arrive at the same opinion. If there is one person in the group with input zero, then all members will eventually agree with his initial opinion, for he will influence the others but no one will influence him.

Theorem 4. In a weakly connected group the members will not reach common agreement except under special conditions in the distribution of initial opinions.

A (strictly) weak digraph contains at least one pair of points with no directed path between them. Thus there are at least two members who cannot influence each other either directly or indirectly.

The disconnected group. When the final equilibrium has been reached, a disconnected group will tend to have at least as many different opinions as there are cliques (i.e., disjoint classes of members), because no clique can influence any other. If all the cliques are themselves either completely connected, strongly connected, or unilaterally connected, it follows from Theorems 1, 2, and 3 that there will be uniformity of final opinions within each clique; but there will be differences among them except under special conditions in the distribution of initial opinion.

Summarizing the theorems illustrated in Fig. 2, we can say that there is a "funnelling effect," a tendency for the opinions of individuals to converge toward one another, and the strength of this tendency increases with increasing connectedness in the power structure of the group. Since the power structure includes the special case of the attraction power of the group, we have a more general group of theorems consistent with the finding that the strength of group standards is determined by the cohesiveness of the group. Additionally the model predicts the exact level of the group standard as well as the precise degree of conformity at each unit. Thus we have rigorously derived a more differentiated statement of the empirically well-established relation between cohesiveness and group standards.

So far we have considered only all-or-none variations in the power of A over B; now we will illustrate the effect of continuous variation.

Theorem 5. The greater the bases of power of A over B (B's attraction to A, B's acceptance of A as an expert, etc.), the more influence A will have on B and subsequently on any other person P for whom there exists a directed path from B to P.

According to Postulates 1 and 3, increases in the basis of power of A over B will increase the strength of the resultant force exerted by A on B and therefore the amount of change produced in B. Similarly in subsequent units this influence will be transmitted, though in a weakened form, from B to P.

The Effects of Communication Patterns

In the preceding section we have dealt with the restricted case of groups of persons whose power is always utilized in every unit.[5] Earlier we noted that the head of an organization may not communicate to all those over whom he has direct power but will instead follow the established channels of communication. Likewise in a face-to-face group a member may remain silent or may attempt to influence some but not

[5] It is probable that B will respond partly to the *relationships* among successive influence attempts by A, for example to the consistency among his various arguments or to the simple fact of too much reiteration of the same influence attempt; these factors are omitted from the present model.

others over whom he has power. These patterns of interaction often become stabilized so that they may be treated as more or less consistent channels of communication. It is also clear that the strength of influence attempted can vary continuously, but we shall here treat the communication from A to B as an all-or-none variable so that we can utilize digraph theory.

Now if we reverse the conditions of Theorems 1 through 4 and consider only completely connected power structures with variations in the degree of connectedness of the communication channels, we can apply the same four theorems and proofs. For example:

Theorem 1a. For all possible patterns of initial opinion, in a completely connected communication network, the opinions of all members will converge to a common equilibrium level equal to the arithmetic mean of the initial opinions of all the members; and this final common opinion will be reached in one unit.

Similarly, theorems analogous to 2, 3, and 4 can be stated for strong, unilateral, and weak communication networks, respectively. All possible networks in experiments of the Bavelas type (3) are included in these theorems.

Even where stable communication channels do not exist, this model may be applied provided the interaction pattern is specified for each unit. Consider a strongly connected cycle of three persons. Theorem 2 states that opinions in this group will converge to a final common opinion equal to $\frac{1}{3}(a + b + c)$. In Theorem 6 we assume a particular communication pattern: A exerts influence in the first unit, B and C exert influence in the second unit, A exerts influence in the third unit, B and C in the fourth unit, and so on.

Theorem 6. In a group where the power structure is a three-person cycle in which A has power over B, B has power over C, and C has power over A, and the communication pattern is A, BC, A, BC, . . . , the final common opinion in the group equals $\frac{1}{3}(2a + b + 2c)$.

We note that a change in the interaction pattern changes the outcome considerably. Furthermore B and C no longer have equal influence, even though they have equal interaction patterns and similar positions in the power structure; it is the interaction of these two factors which produces the difference. A has more influence than B because he comes first in the sequence of interaction, but C has more influence than B because he has direct power over A whereas B's power over A is indirect. Intuitively it would appear that the "primacy effect" shown in this theorem can be generalized: the sooner a person speaks the more influence he will have.

The Effects of Patterns of Opinion

In an experiment like Sherif's, each member communicates to every other and the members probably have relatively equal power. In such a completely connected power structure with completely connected communication channels, what happens to the opinion of a single deviate member?

Theorem 7. The amount of change of the deviate toward the opinions of the majority is proportional to the sum of the deviations of all other members from the deviate.

By Theorem 1 the amount of change by the deviate D equals $d_1 - d$ which is equal to $1/N(d + a + b + c + \ldots) - d$. Thus the more members in the group the more they will influence the deviate. Also the larger each deviation, the more D will change. Though these predictions are generally congruent with Asch's findings, they probably do not agree in detail (1, 15). However, the conditions of Asch's experiment do not fit the model very well.

Leadership

To a large extent leadership consists of a member's ability to influence others both directly and indirectly by virtue of his position in the power structure, including the structure of legitimate authority. Thus leadership may be distributed among many members or concentrated in a few; the pattern of leadership is a distribution which describes the whole group rather than an attribute of single individuals. Figure 2 illustrates the dependence of influence on the total structure of the group.

Compare the influence of member A in the weakly connected group with the influence of member A in the unilaterally connected group. Both groups start out with the same distribution of opinion, and in both groups A has direct influence over only B. However, A's influence is markedly different in the two cases; in the weakly connected group the opinions of others diverge more and more from his, whereas in the unilaterally connected group the opinion of all other members converges completely to A's opinion.

The complete distribution of direct plus indirect leadership in a group with any power structure and any communication network may be calculated by matrix multiplication.[6] We may represent the power structure of the group as a matrix where each row shows the power applied to a member and each column shows the power exerted by a member. A *zero* in the cell corresponding to the ath row and the bth column shows that B does not have power over A, whereas a *one* in the cell corresponding to the cth row and the dth column shows that D does have power over C. Thus the number in a cell represents the number of directed lines from the person in that column to the person in that row (under the conditions assumed in this paper, always *one* or *zero*). If this matrix M is multiplied by itself, then the resulting squared matrix M^2 shows in each cell the number of sequences consisting of two directed lines between the person in the column and the person in the row. The cubed matrix M^3 shows the number of three-line sequences between each pair of persons. By raising the matrix to successively higher powers, we can thus determine the number of directed line sequences, of various lengths, from each member to every other. The matrix M gives the directed lines which will result in influence in the first unit; M^2 gives the two-line sequences through which influence will be exerted by the end of the second unit; M^3 gives the three-line sequences through which influence will be exerted by the end of the third unit; etc.

In order to apply this process to Group G in Fig. 3, we construct a matrix of opinion M where the columns a, b, c, d represent influence exerted by the initial opinions a, b, c, d of persons A, B, C, D, respectively. The rows represent the influence received by these opinions from all the opinions in the group. Thus the cell entries must show the amount by which an opinion is changed by another opinion during one unit; and these values are given by the coefficients in the right hand side of the general difference equations. For Group G these equations are:

$$a_n = a_{n-1}, \tag{1}$$

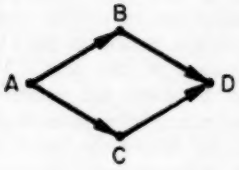
$$b_n = \tfrac{1}{2}(a_{n-1} + b_{n-1}), \tag{2}$$

$$c_n = \tfrac{1}{2}(a_{n-1} + c_{n-1}), \tag{3}$$

$$d_n = \tfrac{1}{3}(b_{n-1} + c_{n-1} + d_{n-1}). \tag{4}$$

The cell a, a has an entry of 1, indicating that in any unit A's opinion is completely determined by his previous opinion; accordingly the remaining cells in

[6] See Harary and Norman (16) for a brief review of some related applications of matrix algebra to sociometric data.



M

	a	b	c	d
a	1	0	0	0
b	$\frac{1}{2}$	$\frac{1}{2}$	0	0
c	$\frac{1}{2}$	0	$\frac{1}{2}$	0
d	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
	2	$\frac{5}{6}$	$\frac{5}{6}$	$\frac{1}{3}$

M^2

	a	b	c	d
a	1	0	0	0
b	$\frac{3}{4}$	$\frac{1}{4}$	0	0
c	$\frac{3}{4}$	0	$\frac{1}{4}$	0
d	$\frac{1}{3}$	$\frac{5}{18}$	$\frac{5}{18}$	$\frac{1}{9}$
	$2\frac{5}{6}$	$\frac{19}{36}$	$\frac{19}{36}$	$\frac{1}{9}$

M^3

	a	b	c	d
a	1	0	0	0
b	$\frac{7}{8}$	$\frac{1}{8}$	0	0
c	$\frac{7}{8}$	0	$\frac{1}{8}$	0
d	$\frac{11}{18}$	$\frac{19}{108}$	$\frac{19}{108}$	$\frac{1}{27}$
	$3\frac{13}{36}$	$\frac{65}{216}$	$\frac{65}{216}$	$\frac{1}{27}$

FIG. 3. The distribution of leadership in a weakly connected group.

row *a* have entries of zero, showing that opinions *b*, *c*, *d* do not influence *a*, since there are no directed paths from *B*, *C*, or *D* to *A*. Cells *b*, *a*, and *b*, *b* have entries of $\frac{1}{2}$ because b_n is a compromise half-way between the previous opinions of *A* and *B*, etc.

It will be noted that each row in *M* (and in M^2 and M^3) sums to 1 because it represents the total opinion of a member, and the fractions along the row represent the proportion of that opinion determined by each person. The sum of a column in *M*, on the other hand, represents the total influence of a person's opinion during the first unit on the opinions of all members (including the influence of his initial opinion on his sec-

ond opinion—which we have called "resistance").

In M^2 the column sum shows the total influence of a person's initial opinion at the end of the second unit (including the changes produced in both the first and second units). Similarly M^3 shows the cumulative influence at the end of the third unit. The same procedure can obviously be extended to any number of units.

Thus the column totals of the successive powers of M give the distribution of leadership over time, as predicted by this theory. In Group G, we can see that A, the only member with input zero in the power structure, continuously increases his influence at the expense of the other members. B and C, having symmetrical positions in the structure, show the same curves of decreasing influence; but D, who is influenced by all other members, has the least influence.

CONCLUSIONS

This theory illustrates a way by which many complex phenomena about groups can be deduced from a few simple postulates about interpersonal relations. By the application of digraph theory we are able to treat in detail the *patterns of relations* whose importance has long been noted by the field theorists. Even if this treatment does not turn out to be empirically correct, it illustrates the need for some such conceptual and mathematical tools if we are to make progress toward the theoretical integration of psychology and sociology.

REFERENCES

1. ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.
2. BACK, K. W. Influence through social communication. *J. abnorm. soc. Psychol.*, 1951, **46**, 9–23.
3. BAVELAS, A. Communication patterns in task-oriented groups. In D. Cartwright & A. Zander (Eds.), *Group dynamics: research and theory.* Evanston: Row, Peterson, 1953.
4. BERKOWITZ, L. Group standards, cohesiveness, and productivity. *Hum. Relat.*, 1955, **7**, 509–519.
5. BIDDLE, B. J., FRENCH, J. R. P., JR., & MOORE, J. W. Some aspects of leadership in the small work group. *USAF Tech. Rep.* (Contract 33 [038]—14091), 1953.
6. CARTWRIGHT, D., *et al. Studies in social power.* Ann Arbor: Institute for Social Research, in press.
7. CARTWRIGHT, D., & ZANDER, A. *Group dynamics: research and theory.* Evanston: Row, Peterson, 1953.
8. CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychologist*, 1955, **10**, 191–198.
9. FESTINGER, L., SCHACHTER, S., & BACK, K. *Social pressures in informal groups.* New York: Harper, 1950.
10. FESTINGER, L., & THIBAUT, J. Interpersonal communication in small groups. *J. abnorm. soc. Psychol.*, 1951, **46**, 92–99.
11. FESTINGER, L., GERARD, H. B., HYMOVITCH, B., KELLEY, H. H., & RAVEN, B. The influence process in the presence of extreme deviates. *Hum. Relat.*, 1952, **5**, 327–346.
12. FRENCH, J. R. P., JR., & GYR, J. Influence and role conformity. In D. Cartwright *et al., Studies in social power.* Ann Arbor: Institute for Social Research, in press.
13. FRENCH, J. R. P., JR., & SNYDER, R. Leadership and interpersonal power. In D. Cartwright *et al., Studies in social power.* Ann Arbor: Institute for Social Research, in press.
14. GOLDBERG, S. Three situational determinants of conformity to social norms. *J. abnorm. soc. Psychol.*, 1954, **49**, 325–329.
15. GUETZKOW, H. (Ed.) *Groups, leadership and men.* Pittsburgh: Carnegie Press, 1951.
16. HARARY, F., & NORMAN, R. Z. *Graph theory as a mathematical model in the social sciences.* Ann Arbor: Institute for Social Research, 1953.
17. HARARY, F., NORMAN, R. Z., & CARTWRIGHT, D. *Introduction to digraph theory for social scientists.* Ann Arbor: Institute for Social Research, in press.
18. HEIDER, F. Attitudes and cognitive organization. *J. Psychol.*, 1946, **21**, 107–112.
19. HOVLAND, C., & WEISS, W. The influence of source credibility on communication

effectiveness. *Publ. Opin. Quart.*, 1952, **15**, 635–650.

20. KATONA, G. Rational behavior and economic behavior. *Psychol. Rev.*, 1953, **60**, 307–318.
21. KELMAN, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. abnorm. soc. Psychol.*, 1950, **45**, 267–285.
22. LEWIN, K. The conceptual representation and measurement of psychological forces. *Contr. psychol. Theory*, 1938, **1**, 1–247.
23. LEWIN, K. *Field theory in social science.* New York: Harper, 1951.
24. LIBO, L. *The cohesiveness of groups.* Ann Arbor: Institute for Social Research, 1953.
25. LIPPITT, R., POLANSKY, N., & ROSEN, S. The dynamics of power. *Hum. Relat.*, 1952, **5**, 37–64.
26. MAUSNER, B. The effect of prior reinforcement on the interaction of observer pairs. *J. abnorm. soc. Psychol.*, 1954, **49**, 65–68.
27. MAUSNER, B. The effect of one partner's success or failure in a relevant task on the interaction of observer pairs. *J. abnorm. soc. Psychol.*, 1954, **49**, 557–560.
28. MOORE, H. T. The comparative influence of majority and expert opinion. *Amer. J. Psychol.*, 1921, **32**, 16–20.
29. NEWCOMB, T. M. An approach to the study of communicative acts. *Psychol. Rev.*, 1953, **60**, 393–404.
30. SCHACHTER, S. Deviation, rejection, and communication. *J. abnorm. soc. Psychol.*, 1951, **46**, 190–207.
31. SCHACHTER, S., ELLERTSON, N., MCBRIDE, DOROTHY, & GREGORY, DORIS. An experimental study of cohesiveness and productivity. *Hum. Relat.*, 1951, **4**, 229–238.
32. SHERIF, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, **27**, No. 187.

(Received March 23, 1955)

Psychological Review
Vol. 63, No. 3, 1956

DRIVE, INCENTIVE, AND REINFORCEMENT [1]

JOHN P. SEWARD

University of California, Los Angeles

Why does one generally do what leads to pleasure and avoid doing what leads to pain? To many persons this is like asking: What color is that white cat? A few, however, may reply: Is that really what one does? It is to these doubters—psychologists, if you will—that the present article is addressed.

Let us start by granting much of the hedonistic premise: human beings do tend, on the whole, to choose activities with pleasant consequences and to abandon those with painful ones. The subjective terminology need not deter us. Thorndike (45) has already told us how to define "satisfiers" and "annoyers" in behavioral terms. His original law of effect—that the connection between a situation and a response is strengthened by satisfaction and weakened by discomfort—could be stated quite objectively. Moreover, if by "connection" he had meant simply the probability that response R would occur in situation S, he need never have discarded one-half of the original law. So interpreted, the law of effect becomes a behavioral statement of hedonism. As such it can be extended to a wide range of species.

But having conceded this much we find ourselves not far ahead. To put the law to use we must know what conditions make for satisfaction and discomfort and how behavior varies with these and other conditions. In a word, we must know *how* rewards and punishments exert their influence. Only by deriving the law itself can we understand apparent exceptions to it. Hedonism in the guise of the "empirical law of effect" (10) may "explain" the obedient child; more basic principles are needed to explain the "problem child" as well.

We face anew, therefore, the question of how rewarded responses are selected, or *reinforced,* and punished responses eliminated. (*Reinforcement* is here used in a strictly empirical sense to refer to the increase in probability of a response accompanied by a reward.) I propose to discuss two theories of reinforcement in the light of experimental evidence, to show that these theories lead to a dilemma, and to suggest a solution.

Two Theories of Reinforcement

Drive-Reduction Theory

The first theory is that of drive reduction, most closely identified with Hull. (By *drive* I mean an excitatory state produced by a homeostatic disturbance.) The primary reinforcing agent, for Hull, was "the diminution in the receptor discharge characteristic of a need" (20, p. 80); [2] when it occurred near the conjunction of a stimulus and a response it produced an increment of habit strength between them. Hull's associates disagree in their use of motivational constructs. Brown (5) and Farber (14) assign the reinforcing function to reduction of a general energiz-

[1] The gist of this paper was presented in a panel discussion: "How can behavior theory best handle the construct of motivation?", Section I, American Association for the Advancement of Science, Berkeley, Calif., December 28, 1954.

[2] In his final volume Hull changed this definition to "the rapid diminution in the motivational stimulus (S_D or s_0)" (2, p. 6). By including s_0 Hull revealed a significant trend in his thinking which he never made fully explicit; had he done so he might have anticipated much of the present discussion.

ing component, *drive*, restricting the drive stimulus to an associative and discriminatory role. Dollard and Miller (13), with Hull, give the drive stimulus a twofold function as cue and reinforcer. But in any case rewards are held to select responses through the learning that results from diminution in an excitatory state.

The scope of drive-reduction theory is vastly enlarged by the assumption that both drives and rewards can be learned. This assumption provides two possible mechanisms of secondary reinforcement:

1. The first, proposed by Mowrer (28) to account for avoidance learning, involves the reduction of a learned drive. If a stimulus has been conditioned to a drive-producing response, removal of that stimulus should serve as a reinforcer. Thus a buzzer that has preceded a shock will arouse fear; turning off the buzzer when a certain response is made should then increase the probability of that response. This type of adjustment has been repeatedly verified (26).

2. The other mechanism of secondary reinforcement is the one postulated by Hull (20). In comparison with the first it may be characterized as the learned reduction of a drive. If a stimulus is conditioned to the response of relaxation that occurs when a need is gratified, it will arouse some part of that response in the absence of the original reward. Since learning often occurs in situations one or more steps removed from primary need gratification, the value of such an explanatory device is obvious. It is worth noting that this mechanism, contrary to the first, has been most frequently verified in connection with appetites rather than aversions (26).

Although—or perhaps because—constantly bombarded by opposing views, drive-reduction theory has displayed remarkable resilience, surviving many an "ordeal by experiment." In the last few years a new rival has arisen to challenge its validity.

Incentive Theory

Sheffield's (39, 41) theory emphasizes the role of the consummatory response in reinforcement. Developed in close correlation with experimental evidence, his view turns out to have much in common with my concept of tertiary, or goal-produced, motivation (37). Briefly his argument may be illustrated as follows:

A hungry animal in a runway enters a goalbox and eats food. On later trials the eating response is generalized to the runway and most vigorously activated when the animal approaches the goal. Since eating cannot be completed without food, however, its partial blocking generates a kind of frustration drive, which energizes whatever response is in progress. The more vigorous the consummatory response, the stronger will be the drive produced by its partial arousal. It follows that the response sequence leading most directly to the goal will induce the most intense drive and receive in turn the most potent excitation. For this reason Sheffield suggested that his account be called the *drive-induction theory* of rewards to contrast it with the older view. To avoid confusing two sources of drive, however, I prefer to call his an incentive theory.

EXPERIMENTAL EVIDENCE

Space limits forbid a detailed review of relevant experiments. In this section we are mainly concerned with the logic of inquiry and the interpretation of results.

The responses involved in satisfying a bodily need may be arbitrarily marked off in three overlapping but roughly sequential stages: (a) instrumental and preparatory responses (R_i), highly vari-

able and largely learned; (*b*) consummatory or goal responses (R_G)—e.g., chewing, swallowing, copulating—relatively uniform within a species; (*c*) homeostatic processes (G)—e.g., digestion, ejaculation—by which the disturbed balance is temporarily restored. The two theories under discussion were advanced to explain the selection of R_t: one does so by invoking G; the other, R_G. To decide between them means must be devised to break up the natural sequence so as to separate stages (*b*) and (*c*).

Efforts to this end have taken two directions. One aims to preserve R_G but prevent G; e.g., by sham feeding. If under these conditions R_t increases in strength or resists extinction, we may infer that drive reduction is not necessary to reinforcement. If, on the other hand, R_t remains weak or extinguishes, it follows that R_G is not a sufficient reinforcing condition. The second method is to bypass R_G and induce G directly; e.g., by stomach fistula. By strictly analogous reasoning, if this method reinforces R_t, we conclude that R_G is not a necessary condition; if it does not, drive reduction is shown to be insufficient.

Goal Response Without Drive Reduction

Here we shall consider only sham-feeding experiments using saccharine,[3] a nonnutritive substance relished by rats. Its effectiveness as a reward was demonstrated by Sheffield and Roby's well-known experiment (40), and confirmed by Cockrell (11). Our question is whether to attribute these results to R_G, as incentive theory requires, or to vicarious drive reduction.

Sham feeding may well induce some of the visceral responses, with accompanying relaxation, characteristic of real feeding. Carlson (8) reported that, in his subject with a gastric fistula, tastes, chewing movements, and swallowing inhibited stomach contractions and the accompanying hunger pangs. (He noted that such stimulation frequently augments appetite.) The inhibiting effect may be either reflex or learned; if learned it coincides with the mechanism of secondary reinforcement suggested by Hull. Carlson's observations on dogs and men, however, were not true of rabbits and guinea pigs, more closely related to rats (8, p. 168).

The argument for secondary reinforcement implies several consequences:

1. By stimulus generalization the reward value of saccharine or any other nonnutritive substance should depend on how closely its flavor resembled that of the animal's normal diet. Where saccharine would fall on such a gradient is unknown.[4]

2. The effect should eventually extinguish through differential reinforcement. Sheffield and Roby found no such trend in three experiments covering 42 days of training, though this may not have been long enough.

3. Saccharine should be less effective, in the long run, than sugars. Results conflict, Carper (9) finding the expected difference, and Thomson (42) finding none. From the standpoint of incentive theory, of course, such comparisons are meaningless until stimulus concentrations are established for which nonhungry rats show no taste preference.

4. If pretraining is given to associate a sweet taste with need reduction, its

[3] Exploratory experiments by Hull *et al.* (22) on one dog with an esophageal fistula gave promise that pursuit of this method should yield rich returns.

[4] Measuring nerve impulses from the rat's tongue, Beidler found that responses to saccharine resembled those to salts and acids rather than sugars; moreover, "a single fiber responding to salt stimulation also responds to saccharin but not to glucose or sucrose" (2, p. 598).

reward value should be enhanced. One test of this prediction (41) gave negative results.

In summary, it appears that R_G's can reinforce behavior even when they fail to satisfy the relevant need. Whether this ability is inherent or merely borrowed as learned drive reduction cannot yet be decided.[5]

Drive Reduction Without Goal Response

Escape and avoidance learning. To test the incentive hypothesis—that R_G can reinforce R_t in the absence of G—it was necessary to select an appetite, such as hunger or sex, with a characteristic R_G and artificially prevent its normal consequence. Drive-reduction theory, on the other hand, can be most readily verified by using noxious external stimuli, since the drives they produce (pain, fear, annoyance) can be manipulated independently of innate R_G's. Organisms are equipped, it is true, with flexion, startle, eyelid, and other defensive reflexes. But no one of these can be relied on to eliminate a disturbance (e.g., a buzzing fly), and other responses are typically brought into play. A rat shocked on a charged grid has a variety of responses available and he will learn to make whichever one consistently lessens or turns off the shock or shock signal, whether this is to sit or press a pedal (29), to jump in the air, or run (31). Such demonstrations point to a large and important class of reinforcers to which incentive theory, as here defined, does not apply.

Other methods. Drive-reduction theorists, on the other hand, do not confine themselves to the aversions. While Sheffield and his coworkers gathered evidence for the importance of R_G, Miller and Kessen (27) in the same laboratory were reinforcing R_t by pouring food directly into a rat's stomach by fistula. True, mouth-fed rats learned much faster. But on the basis of other experiments (3, 24) the authors claimed that hunger was better satisfied by mouth.

Earlier experiments (18, 34) indicated that rats deficient in B vitamins acquired a preference for foods containing them, apparently due to beneficial aftereffects. Thomson and Porter (44) recently reported learning in sodium-deprived rats rewarded by a subliminal salt solution. But probably the purest test of a need-reduction theory was provided by Coppock and Chambers (12), who taught hungry rats a head movement by rewarding them with an intravenous injection of glucose.

Evidently drive reduction can select responses in the absence of any characteristic goal response or goal stimulus.

Interpretation

These experiments leave us "on the fence." There is evidence for both consummatory and homeostatic factors in reinforcement. The chief unsettled question is to what extent R_G owes its potency to previous association with G. This question, however, leads to another that may help us to answer the first.

If R_G is merely a secondary reinforcer of type (b), its antedating component (r_G) should also reduce the appropriate drive to some extent. But Sheffield's view, we recall, is just the opposite: as a frustrated response, r_G produces an added drive stimulus. Does r_G function as drive or as partial reward? A final answer has not yet been found (38), but some evidence for Sheffield's view is already at hand (1, 15, 16, 33, 39).

[5] The experiments of Sheffield, Wulff, and Backer (42) and Kagan (23), using "sham coitus," should also be read with this problem in mind.

A CHOICE-POINT DILEMMA

Since the two theories are not strictly incompatible, it is possible to accept both. Indeed the foregoing evidence strongly indicates that both drive reduction and incentive play a part in reinforcement. But this conclusion resolves one dilemma only to trap us in another.

I have already alluded to this new dilemma in another context (36). Consider rat A at the choice point of a T maze. The entire floor of the maze is a charged grid except for the right-hand goal box. There is little doubt that in the course of repeated noncorrection trials the average rat will develop a strong right-turning habit. Followers of both Hull and Guthrie will agree, but for different reasons, that the key to this outcome is the removal of the drive stimulus, shock, in the goal box.

Now consider rat B at a similar choice point but without shock. Instead, the rat is hungry and the entire maze is empty except for the right-hand goal box, which contains a few drops of saccharine solution. Again, as Sheffield and Roby have shown, the rat will learn to turn right. Suppose we reject the tenuous assumption that the taste of saccharine has the power, innate or acquired, to reduce hunger, but grant it the power to evoke drinking. Resorting to Sheffield's incentive hypothesis, we infer that right-turning stimuli arouse incipient drinking and that this frustrated R_G produces more drive to strengthen the ongoing act.

At this point we confront the dilemma: How can rat A choose the path of less drive while rat B chooses the path of more drive? If shock and appetizing "food" are both sources of drive to an organism at a choice point, how can he learn to escape or avoid one and approach the other? I contend that these are not exceptional cases but typical of aversions and appetites in general. If so we face a problem of fundamental significance.

Brown (6) recently discussed in this Journal a related question. How, he asked, can drive-reduction theory accommodate the fact that responses are sometimes strengthened in the face of increasing stimulation? Though much of his paper was devoted to distinguishing between stimulation and drive, he considered the special case in which a response is accompanied or followed by drive increase without disruption. His answer, closely allied to the one to be offered below, was that a drive will intensify an ongoing act if the drive stimulus is not connected with interfering responses. But for a drive-reduction theorist our choice-point dilemma does not arise, since for him reward training does not involve a drive increment. Only in the light of incentive theory and the evidence for it does this problem become unavoidable. For this reason Brown could not raise the further question: how does his principle—that drive stimuli disrupt behavior only if associated with conflicting reactions—enable us to predict opposite outcomes for escape and reward training?

My own attempts to solve this problem, starting five or ten years ago, reflect its stubbornness. In retrospect I find I have vacillated between the two viewpoints here brought face to face—between considering rewards as satisfiers and as incentives. A brief review may clarify my present position.

My first reaction to the dilemma was to assume that increasing drive, or tension, inhibits activity in progress while decreasing tension promotes it. On the surface, at least, this seems a plausible premise: by equating "tension-reducing" with "pleasant" and "tension-building" with "unpleasant" we can say with the hedonists—and neo-hedonists like McClelland (25) and Young (46)—that

organisms act in such a manner as to maximize pleasure and minimize unpleasantness.

On this assumption I introduced the concept of *tertiary motivation* (35) to suggest that secondary reinforcers act by "realizing" a goal set (rs_G), thereby lowering its tension. Mowrer pointed out the error of confusing drive-reducing function with motivation. Later (37) I corrected it by identifying tertiary drive with the *blocking* of rs_G, but still held that drive (rs_D) and goal surrogate were negatively related; release of rs_G diminished rs_D and facilitated responses in progress.

This position now appears untenable. Osgood (32, p. 443), discussing the dual role of the anticipatory goal reaction as both drive and reward, pointed out the difficulty; so did Mowrer (30) in his most recent theoretical statement. To explain delayed reward learning, for example, we should have to argue that approach to the goal permits rs_G to approach completion, promoting the correct response by reducing tertiary drive. But if this were true, drive would be progressively lowered as one neared the goal; the closer one came, the less energy would be available for approach behavior. Unfortunately I can no longer shut my eyes to the evidence for a goal gradient both in strength of approach (4, 19) and in height of tension (16, 39).

The alternative is to start from the opposite premise: viz., that rewards function as incentives. Once before (36) I explored some of the possibilities of this view. A motive was described in terms of two moments, *drive* (rs_D) and *goal* (rs_G). This much I believe was on solid ground. To account for the goal direction of behavior, however, I simply weighted rs_G more heavily than rs_D in a formula for motive strength. This step was probably abortive. Even if a method could be devised for empirically determining the weights, the formula was *ad hoc*, with little rationale beyond the observations it was designed to explain.

Whether this admittedly trial-and-error process has helped or hindered, I believe the materials are now available for a more fruitful approach.

TOWARD A SOLUTION

From previous treatments we salvage two motivational constructs, to be classed as mediating or response-produced stimuli, one concerned with instigators, the other with satisfiers. The first is the *surrogate drive stimulus* or *drive* (rs_D). It is produced by certain disturbers of homeostasis, whether external (e.g., pain and pain signals) or internal (e.g., metabolic deficit); it is reduced by removal of the disturbing stimulus or by replacement of the deficit. Drive has two functions, associative and nonassociative: as *motor* it energizes responses in general; as *cue* it becomes conditioned to specific responses. These two functions correspond roughly to D and S_D in Hull's system.

For the second construct the terms *surrogate goal stimulus*, *tertiary drive*, and *goal* (rs_g) are interchangeable. It is produced by the strong incipient arousal of a response under conditions that prevent its completion. Any such response qualifies as a *goal response* (R_g): one large class undoubtedly consists of consummatory responses (R_G); another and more inclusive class, of responses coincident with drive removal. In contrast with drives, goals may be reduced in two ways: by removing the stimulus to R_g and by completing R_g. In common with drives, goals have the same two functions. A fractional R_g produces cues that may be conditioned to particular responses; attempts to complete it ahead of time generate nonspecific *goal tension*. The closest analo-

gies in current theory are Brown and Farber's (7) *frustration* constructs F and S_F, although rs_g is used in situations these authors would hardly call frustrating.

We now ask what are the roles of rs_D and rs_g in response selection? My answer is simple and far from original. Paraphrasing Guthrie (17), I suggest that both drive and goal surrogates, along with the external stimuli giving rise to them, become conditioned to each response made in their presence and remain so until conditioned to an incompatible response.

To apply this principle to response selection we need an estimate of the strength of motivational stimuli that will be and stay conditioned to the R_i in question. On the assumption that any such stimulus remaining after R_i has been made will be unconditioned, we may introduce the concept of *effective stimulus intensity* ($\bar{S}$), defined as the difference between intensity before (S_B) and intensity after (S_A) R_i. For *effective drive surrogate* we may write:

$$\overline{rs}_D = rs_{DB} - rs_{DA}.$$

Similarly, *effective goal surrogate* may be written:

$$\overline{rs}_g = rs_{gB} - rs_{gA}.$$

Note, however, that in practice $\overline{rs}_g$ can often be more readily computed from some measure related to R_g itself, such as amount of reward.

Let us now return to the choice point and try to predict the behavior of rats A and B. To simplify the task we shall assume other things equal and evaluate only the part played by motivational cues. First, how does rat A learn to escape the charged grid by taking the right-hand path? In the absence of a definite R_G drive stimuli presumably play the dominant role. After a number of trials the $\overline{rs}_D$'s conditioned to the left and right paths will approximate the following values:

$$\text{Left: } rs_{DB} - rs_{DA} = 0,$$
$$\text{Right: } rs_{DB} - 0 = rs_{DB}.$$

Rat A will therefore turn right.

Secondly, how does rat B learn to take the right-hand path to saccharine? Hunger-drive stimuli are presumably diminished only slightly, if at all, so they contribute little to *selective* motivation. After a number of trials, however, a goal surrogate is aroused by the right-path stimuli and becomes conditioned to an approach response. It remains conditioned, moreover, because it is removed by the actual drinking of saccharine in the goal box. Since no rs_g is conditioned to left-hand stimuli or responses except by generalization, rat B will turn right.

Even for a simplified version the above account is oversimplified. It is highly probable that in the course of repeated escapes from shock rat A will learn some response to the goal box that will generalize to the choice point as rs_g. This process should then contribute to total reinforcement by the mechanism already described for saccharine. Similarly, in training to a nutritive reward if not to saccharine itself, rat B's hunger should act just like rat A's shock in strengthening the correct response. How drive and goal combine and interact is a problem for future development.

This is only a fragment of theory. It needs to be fitted into a more comprehensive framework before its implications can be tested or even derived. I believe this can be done and have taken some steps toward modifying Hull's equations in order to accommodate the proposed concepts. Enough of the theory is given, however, to reveal its roots and one or two branches. It is founded squarely on Guthrie's doctrine of association by contiguity, extending his principle to a special class of stimuli, mo-

tives, with two subclasses, drives and goals.[6] It contains more than a hint of Tolman's expectancy in its notion of goal surrogate, and its hope of elaboration owes much to Hull. This eclecticism calls for no apology. Current theorizing, with its increasing reliance on mathematical tools, displays a healthy tendency to disregard earlier lines of cleavage.

SUMMARY

Scratch a motivation theory and you find the problem of reinforcement. This paper deals with two possible reinforcers: drive reduction and consummatory response. Experiments designed to separate their influence have yielded evidence in support of both. Drive reduction seems to be a sufficient condition for response selection. But rewards apparently do more than reduce drives; as incentives they appear to heighten excitement as well as activate instrumental responses. This poses a problem: How can drive reduction and incentive *induction* both produce the same result, i.e., strengthen a concurrent response? It is suggested that Guthrie's theory, with the aid of two motivational constructs, *drive* and *goal*, offers a solution.

REFERENCES

1. AMSEL, A., & ROUSSEL, JACQUELINE. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.
2. BEIDLER, L. M. Properties of chemoreceptors of tongue of rat. *J. Neurophysiol.*, 1953, **16**, 595–607.
3. BERKUN, M. M., KESSEN, MARION L., & MILLER, N. E. Hunger-reducing effects of food by stomach fistula versus food by mouth measured by a consummatory response. *J. comp. physiol. Psychol.*, 1952, **45**, 550–554.
4. BROWN, J. S. Gradients of approach and avoidance responses and their relation to level of motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.
5. BROWN, J. S. Problems presented by the concept of acquired drives. In *Current theory and research in motivation: a symposium*. Lincoln: Univer. of Nebraska Press, 1953. Pp. 1–21.
6. BROWN, J. S. Pleasure-seeking behavior and the drive-reduction hypothesis. *Psychol. Rev.*, 1955, **62**, 169–179.
7. BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.
8. CARLSON, A. J. *The control of hunger in health and disease*. Chicago: Univer. of Chicago Press, 1916.
9. CARPER, J. W. A comparison of the reinforcing value of a nutritive and a non-nutritive substance under conditions of specific and general hunger. *Amer. J. Psychol.*, 1953, **66**, 270–277.
10. CARR, H. A. The law of effect. A round table discussion. I. *Psychol. Rev.*, 1938, **45**, 191–199.
11. COCKRELL, J. T. Operant behavior of white rats in relation to the concentration of a non-nutritive sweet substance used as reinforcement. Unpublished doctor's dissertation, Indiana Univer., 1952.
12. COPPOCK, H. W., & CHAMBERS, R. M. Reinforcement of position preference by automatic intravenous injections of glucose. *J. comp. physiol. Psychol.*, 1954, **47**, 355–358.
13. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
14. FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska symposium on motivation 1954*. Lincoln: Univer. of Nebraska Press, 1954. Pp. 1–46.
15. GEIER, F. M. The measurement of tension in the rat. *J. comp. Psychol.*, 1942, **34**, 43–49.
16. GEIER, F. M., & TOLMAN, E. C. Goal distance and restless activity. I. The goal gradient of restless activity. *J. comp. Psychol.*, 1943, **35**, 197–204.
17. GUTHRIE, E. R. *The psychology of learning*. (Rev. Ed.) New York: Harper, 1952.

[6] Readers acquainted with my earlier efforts (see **36**) may find this amusing. It was the problem of reinforcement that first led me away from Guthrie's theory; the same problem, like Maeterlinck's bluebird, now brings me back sadder and wiser (?) to my starting point.

18. HARRIS, L. J., CLAY, J., HARGREAVES, F. J., & WARD, A. Appetite and choice of diet. The ability of the vitamin B deficient rat to discriminate between diets containing and lacking the vitamin. *Proc. roy. Soc. Lond., Ser. B*, 1933, **113**, 161–190.
19. HULL, C. L. The rat's speed-of-locomotion gradient in the approach to food. *J. comp. Psychol.*, 1934, **17**, 393–422.
20. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
21. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
22. HULL, C. L., LIVINGSTON, J. R., ROUSE, R. O., & BARKER, A. N. True, sham, and esophageal feeding as reinforcements. *J. comp. physiol. Psychol.*, 1951, **44**, 236–245.
23. KAGAN, J. Differential reward value of incomplete and complete sexual behavior. *J. comp. physiol. Psychol.*, 1955, **48**, 59–64.
24. KOHN, M. Satiation of hunger from food injected directly into the stomach versus food ingested by mouth. *J. comp. physiol. Psychol.*, 1951, **44**, 412–422.
25. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
26. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 435–472.
27. MILLER, N. E., & KESSEN, MARION L. Reward effects of food via stomach fistula compared with those of food via mouth. *J. comp. physiol. Psychol.*, 1952, **45**, 555–564.
28. MOWRER, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, **46**, 553–564.
29. MOWRER, O. H. An experimental analogue of "regression," with incidental observations on "reaction-formation." *J. abnorm. soc. Psychol.*, 1940, **35**, 56–87.
30. MOWRER, O. H. Two-factor learning theory reconsidered with special reference to secondary reinforcement and the concept of habit. *Psychol. Rev.*, 1956, **63**, 114–128.
31. MOWRER, O. H., & LAMOREAUX, R. R. Fear as an intervening variable in avoidance conditioning. *J. comp. Psychol.*, 1946, **39**, 29–50.
32. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
33. ROUSSEL, JACQUELINE S. Frustration effect as a function of repeated non-reinforcements and as a function of the consistency of reinforcement prior to the introduction of non-reinforcement. Unpublished master's thesis, Tulane Univer., 1952.
34. SCOTT, E. M., & VERNEY, E. L. Self-selection of diet. VI. The nature of appetites for B vitamins. *J. Nutrition*, 1947, **34**, 471–480.
35. SEWARD, J. P. Secondary reinforcement as tertiary motivation: a revision of Hull's revision. *Psychol. Rev.*, 1950, **57**, 362–374.
36. SEWARD, J. P. Introduction to a theory of motivation in learning. *Psychol. Rev.*, 1952, **59**, 405–413.
37. SEWARD, J. P. How are motives learned? *Psychol. Rev.*, 1953, **60**, 99–110.
38. SEWARD, J. P., & PEREBOOM, A. C. Does the activity wheel measure goal striving? *J. comp. physiol. Psychol.*, 1955, **48**, 272–277.
39. SHEFFIELD, F. D., & CAMPBELL, B. A. The role of experience in the "spontaneous" activity of hungry rats. *J. comp. physiol. Psychol.*, 1954, **47**, 97–100.
40. SHEFFIELD, F. D., & ROBY, T. B. Reward value of a non-nutritive sweet taste. *J. comp. physiol. Psychol.*, 1950, **43**, 471–481.
41. SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. comp. physiol. Psychol.*, 1954, **47**, 349–355.
42. SHEFFIELD, F. D., WULFF, J. J., & BACKER, R. Reward value of copulation without sex drive reduction. *J. comp. physiol. Psychol.*, 1951, **44**, 3–8.
43. THOMSON, C. W. Need reduction and primary reinforcement: A comparison of the reinforcing value of sugar and saccharine. Paper read at West. Psychol. Ass., Long Beach, May 20, 1954.
44. THOMSON, C. W., & PORTER, P. B. Need reduction and primary reinforcement: maze learning by sodium-deprived rats for a subthreshold saline reward. *J. comp. physiol. Psychol.*, 1953, **46**, 281–287.
45. THORNDIKE, E. L. *Animal intelligence: experimental studies*. New York: Macmillan, 1911.
46. YOUNG, P. T. The role of hedonic processes in the organization of behavior. *Psychol. Rev.*, 1952, **59**, 249–262.

(Received May 12, 1955)

Psychological Review
Vol. 63, No. 3, 1956

INFORMATION-RECEIVING BEHAVIOR OF MAN

MASANAO TODA

Hokkaido University

The recent development of communication theory has certainly stimulated the corresponding area of psychology, and many experiments have been designed and successfully performed, mostly for the purpose of ensuring the applicability of concepts involved in the theory to psychological analysis. It is no wonder that the theory contributes much to the analysis of human communicative behavior, since at almost any time any person is receiving information from his surrounding world (perceiving), storing it (memorizing), as far as his communicative behavior is and sending it (conducting); and thus, concerned, he can always be regarded as a kind of transducer in a communication system—i.e., as a receiver of power from one system and a supplier of power to a different system.

A psychologist would be at a loss, however, if he wished to go a step further into the detailed complexity of actual human communicative behavior by directly applying the outcomes of the theory, since, though man is surely a kind of transducer, he is one of the types hitherto never discussed. Present-day communication theory has been developed mainly to deal with the problems of engineering communication, of which a distinguished characteristic is that it is carried out under rather stable boundary conditions, serving a rather limited variety of purposes. Thus it is a relatively easy task to find the most efficient acting transducers when certain boundary conditions are given. On the other hand, man performs many other functions than that of transducer and thus, even when communicating, he should serve a rich variety of purposes, many of which may be entirely irrelevant to efficient transmission. For example, as the receiver, he should not only know what the original message is but should extract from signals the information content which is usually called ***meaning;*** as the memorizer, he should select the information content to be memorized according to its degrees of importance; and finally, as the sender, he will sometimes want to send wrong information to deceive his enemy. And all of those performances have usually to be done under rather unstable conditions.

What is unstable in human communication is not man himself. There are many evidences, mostly physiological, which make one suspect that man has surprising capabilities in stabilizing himself as a mechanical system. Most of the causes of the instability of human communication systems are to be looked for within the instability of the ***relations*** between man and his surrounding world. Unlike Shannon's mechanical receiver (4), man usually has to deal not with a single ensemble but with a large variety of them simultaneously, most of which could possibly be statistically unstable. Further, no proper rules of decoding are given by senders, and they have to be discovered by the receiver himself; it seems quite improbable that there could be any transducer acting efficiently enough in such an unstable and complex situation. Certainly, there exist no mechanical transducers as yet which meet the requirement. The only possible type of transducer which could meet it should at least satisfy the following two conditions: it should be capable by itself of developing adequately

efficient rules of decoding separately for different information sources; and it should be capable of changing the rule of decoding as rapidly as possible whenever it receives the information that the statistical structure of an ensemble seems to have changed.

In other words, the competent transducer should be one that learns; for *learning* roughly means those two functions, gathering information from the surrounding world, and changing the performance in a way that suits the environment best. No one would deny that man is one such type of a system, i.e., one that learns, although it is doubtful if he learns most efficiently as a transducer. But, if we take into account the severe conditions under which he should act as a transducer—the facts that he should perform many tasks other than communication, that he should act below a certain rate of energy expenditure, and so forth—we ought to consider him a fairly, though not the most, efficient type of transducer. As far as this is taken for granted, it should be possible, though difficult, to find the principles or a theory formulating the functions of a transducer of this type, or of a man as a transducer. In the following sections we shall outline what considerations are needed for the theory, restricted mostly to the theory of the human *receiver*, and not of the human transducer as a whole.

COMMUNICATION AS A GAME

One of the most distinctive features of human communication, not yet discussed fully, is that it requires game-theoretical considerations. In engineering communication, the purposes of sender and of receiver are usually in perfect congruence; the sender wishes to send as much information as possible and the receiver wishes to know what the original message is as correctly as possible. On the other hand, in actual human communication they are not always in such accord. Sometimes the sender is indifferent to the receiver; signals fall only incidentally on the receiver, or the sender knows nothing about the fact that his emitted signals are caught by the receiver. Sometimes the sender is hostile to the receiver, and he intends to deceive the receiver by sending signals which would convey wrong information if decoded by conventional rules. In the latter case they are in a game situation. Even in the former case, the receiver often has to be careful not to be deceived, and has to behave as if he were in a game.

This is clearly illustrated in A. Wald's theory of statistical decision functions (8), in which it is formalized that a statistician intending to decode signals from Nature (apparently an indifferent sender of information) should sometimes be careful to use the minimax strategy in establishing his decoding rules. Accordingly, it is clear that game theory will play a particularly important role in the theory of human communication. Von Neumann and Morgenstern's theory (7), however, is not directly applicable to our purposes. Here what is required is a theory that is more specifically developed to be able to describe various patterns of actual human behavior—that takes into account, for example, the fact that not all of the participants in communication games are complete strategists. If a participant found that his opponent is no such thing, or if he gets the information that his opponent's behavior is in some way stereotyped, then to take advantage of that fact would usually give him an optimal strategy better than the minimax strategy. As mentioned above, that some sender emits false messages is one of the typical features of human communication, and in this case the receiver could

find certain decoding rules which provide him with correct information if the sender's strategy of deceiving is in some way stereotyped, because a false message is not a random message and it contains information.

Another point we require of game theory is much more serious. The theory of games assumes that the values of objective probabilities in chance moves are known to all participants. This, however, is scarcely realized in the actual situations of games. In poker, for example, where the probability of each hand is objectively determined, the exact values of probability are not known to the participants. If one asks them what the exact values are, even the most skilled players would not be able to tell. Nevertheless, the considerable rationality observable in the strategies of skilled players implies that they have some approximate values of probability, perhaps subconsciously obtained through their experience in the game; the necessary information has undoubtedly been given to them from their previous experience, i.e., from the time series of poker products previously observed. In other words, it is contended that all the values of probabilities with which poker players play are estimated values, and the superiority of skilled players is due, at least partly, to the fact that the numbers of their observed samples exceed those of beginners. The fact that the estimated values of probabilities could be different from participant to participant has such grave influence upon the properties of the game that some strictly determined games could no longer be strictly determined, since the good strategy of a participant ignorant of the true values is, at best, to identify his estimated values with the true ones. This is an essentially important point in human communication.

MULTIPLE INFORMATION EXTRACTION FROM ONE AND THE SAME SIGNAL

The meaning of *meaning* has been one of the most tempting subjects in human communication theory, and a complicated skein of various meanings has been attached to the term. Though it is certainly difficult to unravel the whole skein, a considerable clearing-up nevertheless seems possible by the traditional way of thinking in the communication theory.

Suppose a man hears another's voice making the sounds, "My wife is ill." Then we can say that the signal caught by his ear means that "the words just spoken are 'my wife is ill.'" This is the only one that may be called meaning which is dealt with in Shannon's theory, and hereafter will be referred to as primary meaning. There certainly can be others meant by the same signal, for instance, the semantic meaning that the speaker's wife *is* ill and is neither healthy nor dead. It is obvious that this semantic meaning as well as the primary meaning contains information, since it reduces the uncertainty about the physical conditions of the speaker's wife, just as the primary meaning reduces the uncertainty about the transmitted signal. These two kinds of meaning are both information contents, and hence can be treated in the same theoretical framework. The same argument holds even when a physical or a semantic noise has intervened—e.g., a large crashing sound or a strong foreign accent of the speaker. The semantic meaning in this case is still an information content, though turned by equivocation into a more vague one such as "The speaker's wife is *likely* ill."

It should be noted that the semantic meaning is not the only secondary meaning. The words may also mean that the speaker is then short of money. Very likely this is just what he wished

to mean by the words. Or, perhaps, the signal might be a sign meaning, entirely independently of its semantic context, say, that it is time to start to act. Or, it might mean that the speaker is a liar, if the listener incidentally knows that the wife is quite well. In any case, the meaning is an information content, so far as it indicates this or that among the whole set of possible alternatives.

Ideas introduced in the above paragraphs, with the exception of multiple information extraction, are not entirely new in communication theory. Moreover, multiple information extraction itself is really no more than the consequence of succession of different decodings. Presumably a considerable part of the receiver's functions and operations, which have been discussed by various investigators with different terminologies, could properly be reduced into the following two activities: the selection of a set of alternative states in relation to which information is analyzed, and the establishment of a decoding rule which transforms a particular signal to a particular information content.

These considerations suggest that a further attack on the problem of meaning, using the ordinary language of communication theory, is possible by regarding meaning as a kind of information content, provided that an appropriate measure of the amount of information can be assigned to the contents.

AMOUNT OF INFORMATION AND SUBJECTIVE PROBABILITY

When a psychologist's interest is focused on the human receiver, he should remember that the Shannon-Weaver measure of the amount of information is the measure prepared for the outsider—neither receiver nor sender—who is watching and controlling the communication system, and whose only interest is the mean rate of information transmission, namely, $-f\Sigma_i p_i \log_2 p_i$, where f is the signal frequency. Such a value prepared for or measured by the outside controller is, as a matter of fact, entirely irrelevant to the actual human receiver, who cannot usually control at will the sender's rate of information transmission. Though naturally the receiver will hope that the rate of information transmission will increase, for him the most meaningful measure of the amount of information must be the amount of information intake, which he can vary by changing his decoding rules. The receiver's measure of the amount of information is thus to be the measure of extracted information content, primary or secondary. It is further required that this measure be defined for the individual content and not for the ensemble of contents. In actual human communications the receiver cannot usually get signals frequently enough from the same ensemble to be able to estimate the mean rate of information transmission (or intake) within the time available. If he waited too long, the improvement of his decoding rules would be practically ineffective, or more likely, the statistical properties of the ensemble might change during the time, even though they were once stable.

Such a definition also will be called for by an economic point of view. As Marschak states (1), information contents can be regarded as commodities that can be sold and bought at a certain price. Most information contents are traded in singly and not as an ensemble, and, though the utility of information is certainly a more complex notion than the amount of information, it is natural to suppose that a higher price somehow corresponds to a larger amount of information among the same kind of information content.

At first sight, to define such a measure may seem quite simple. If we let p be the probability of getting a particular information content, then the amount of information of the content is $-log_2 p$ bits. It is obvious that this measure for single content is compatible with Shannon and Weaver's measure for ensemble. Practically, however, this definition is meaningless. Who among ordinary human receivers knows of the objective statistical properties of ensembles? As mentioned previously, we should drop the assumption that all the true probabilities are known by receivers. The values are estimated, and not given from the start. Accordingly, the above definition should be altered, and though the alteration is slight, it would have a far-reaching influence on the concept of information. Let q be the *estimated* value of the probability of getting an information content. Then the amount of information of the content for the receiver is defined as $-log_2 q$ bits. Remember that this is the definition of amount of information and not that of estimated amount of information. I am convinced that this is the only reasonable measure for our purposes, and might be, in a sense, a more fundamental measure than Shannon and Weaver's. The following paragraphs are of central importance to this point.

Consider a simple true-or-false problem which a person does not know how to answer. Its correct answer certainly contains a finite amount of information, since by knowing the answer his uncertainty vanishes and he knows more than he did before. The value of the amount of information is either $-log_2 q$ or $-log_2(1-q)$, depending on whether the correct answer is "true" or "false," where q is his estimated probability that the answer will turn out to be "true." What is the objective probability in this case? If it is to be one or zero, such an answer could contain no information in the classical sense. This difficulty cannot be avoided by taking as the objective probability the frequency that the answer has been "true" in past similar situations. Though his value of probability q might possibly be determined in a somewhat similar manner, no one could find any mathematically rigorous probability of such type, or any statistically stationary ensemble of *similar situations*. The simple true-or-false problem can be replaced by the true-or-false problem of any unproved mathematical statement. As Brouwer once pointed out, an unproved mathematical statement is in a state somewhere between true and false. But who knows or can determine the objective probability that some unproved mathematical statement will turn out to be true? Nevertheless, no one would deny that its proof (whether affirming or negating) contains information, and we have to measure its amount.

Perhaps the reader will admit the well-known definition that information is something which reduces the receiver's uncertainty. Then he would also have to admit the possibility that a message brings different amounts of information to different receivers, since the values of uncertainties could be quite diverse among actual human receivers. For instance, if there are two receivers, one of whom knows of something in its every detail with no remnant of uncertainty and the other has very poor knowledge about it and thus has very great uncertainty, it is easily seen that a message on the subject will give the former nothing and the latter a large amount of information. These facts clearly agree with our definition of amount of information, in which mathematically true probabilities are replaced by estimated probabilities or *subjective probabilities*. Since subjective probability is measurable by some experimental techniques which adequately separate it

from utility (5), the definition stands on a definite operational ground.

THE OBSERVAND AND ITS COMPLETE SET OF STATES

As explained in the preceding section, a human receiver may derive from the same message several information contents, each referring to a different subject, such as, "the words just spoken," "the physical condition of a man's wife," or "his monetary condition." In order to avoid needless ambiguities we shall call the *subject* of an information content the observand, and any information content will be written in the form "An observand X is in a state x_i," or, in the above example, "The wife's physical condition" is in the state "ill." Usually there is more than one state of an observand forming a space to which a subjective probability measure is assigned. Roughly speaking, the set of states is complete when it is exhaustive and the states are mutually exclusive. "Healthy," "ill," and "dead" offer an example of a complete set for the observand "one's physical condition." Of course, there can be other complete sets for the same observand; one's physical condition could be classified by completely different and independent criteria. It seems proper, however, to regard these independent sets as belonging to different observands, since information contents described by each of them can also be independent of each other.

Given any complete set of states, it is possible to create a new complete set that is a relative of the first by simply combining some of the states, or, conversely, by dividing some of the states in the original set. Which one of these related states the receiver decides to use is of great importance, since both the information content and the measure of the amount of information depend on the selection. It is evident that the amount of information will increase as the states are subdivided, but there are limits beyond which further divisions are futile. One obvious limit on how far the receiver should go in differentiating states is reached when differentiation does not produce a more accurate identification of the observand. If a receiver used one of the overdifferentiated sets, he could obtain only a type of information content such as "The observand is in a particular subset including more than one state"; and as to any further precise location of the observand's position among the states in the subset no information could be drawn from the signal. Another restriction may be a physiological one. It is most probable that some principle of economy or efficiency affects the selection of the set, since it is likely that the smaller the number of alternative states the less the energy, the smaller the number of necessary physiological elements, and the less the probability of error in analyzing and memorizing information. Further, we must consider social influences on the selections of the set. In a social group, whose members have assembled to cooperate, communications should be performed most efficiently, and it is quite natural that common sets shared by the group members are established for the observands which are significant for the group activity. Perhaps the social-psychological concepts of group norm and shared frame of reference are relevant to such common sets (2).

GENERALIZATION OF DEFINITIONS

We have already described two types of information content, "X is in x_i" and "X is in S_i," where X, x_i, and S_i signify an observand, its i-th state and its i-th subset of states, respectively. More precisely, however, these contents should be expressed as "X is in x_i (or in S_i)

with probability one." Once expressed in such a form, it is obvious that they are special types of information content. (It should be particularly noted that there is no necessity for coincidence of the forms of verbal message and the information content extracted from it. At least one reason for the fact that verbal messages are apt to take the form of absolute judgments may be found in their relative immunity from noise attack.) Thus, an information content will more generally be expressed in such form as "X is in S_i, S_j, S_k, . . . with probabilities p_i, p_j, p_k, . . . , respectively," where S's are subsets of states. The probabilities are of course subjective probabilities estimated from the receiver's past experience, which might have told him that, when he obtained such signals from such an information source in such a situation, the observand was in such and such states with such and such frequencies; or perhaps it might have told him more. Then we can formulate: An information content is a subjective probability measure assigned to a complete set whose elements are mutually disjunct subsets of the original complete set with which the information is analyzed. If any one of the subsets contains more than one state, the information content will be said to be degenerating in that subset.

Now we have to generalize the definition of amount of information according to the generalization of the definition of information content. This is a rather simple task when it is absolutely certain that the probabilities given by the obtained information content are the objective probabilities. Let the given objective probabilities and the existent subjective probabilities be p_1, p_2, . . . and q_1, q_2, . . . , respectively. Then if it is assumed to be known that X is really in x_i, the amount of information obtained is $-log_2 q_i$ and $-log_2 p_i$ before and after the objective probabilities are given, respectively. Accordingly, the mathematical expectations of amount of information are $-\Sigma_i p_i log_2 q_i$ and $-\Sigma_i p_i log\, p_i$ before and after the probabilities are given, respectively. It is easily verified that the former is not less than the latter. Hence the expectation of amount of information given by the information of objective probabilities is positive and equal to $-\Sigma_i p_i log_2 q_i + \Sigma_i p_i log_2 p_i$. We shall call this quantity the *information gain*, and it coincides with the definition of amount of information previously offered when the given probabilities are one and zero. Even when it is not absolutely certain that the given probabilities are objective, we would use the same equation to estimate the information, as long as we are convinced that the given probability distribution is far more authentic than the original subjective probability distribution. Difficult problems arise when neither one of the distributions has such a predominant authenticity. This problem, however, has to be treated from a more general point of view.

SYNTHESIS OF INFORMATION

It is a common experience to receive from different information sources various information contents about the same observand which more or less contradict each other; this leaves one uncertain as to what is the most reliable information content. This is the problem of synthesis of information contents, which is important not only for individuals but also for social organizations; and our lines of thought seem to provide its rational solution. Because of the considerable complexity of the problem, however, only a brief outline of possible solutions will be shown here, of which the more detailed discussions are found in another paper by the author (6).

Suppose that there are $m - 1$ number of different information contents given to a receiver. Then there are m num-

ber of different probability distributions, including his own subjective probability distribution possessed in advance. Usually each distribution is different in its reliability, according to its source and the situation in which it is given. One quite simple way of obtaining the most reliable probability distribution is to compute weighted means of probabilities for each state, where each weight represents the reliability of the corresponding information content. This idea of simple linear combination can be made more convincing by the following sample case: It is known that some one of the alternative distributions is the objective probability distribution, and the probability that the i-th distribution D_i is objective is W_i, $(i = 1, 2, \ldots, m)$. Let p_{ij} denote the probability given by D_i for j-th state (or subset), and let r_j denote the most reliable probability to be obtained, $(i = 1, \ldots, m$ and $j = 1, \ldots, n$; n is the number of alternative states or subsets). If it happens that D_i is true, then the *expected amount of information* that he loses by adopting the distribution R instead of D_i is $-\Sigma_j p_{ij} log_2 p_{ij} + \Sigma_j p_{ij} log_2 r_j$. We shall call this quantity the *information loss*, which has just the same form as that of the information gain, but with the negative sign, and is never positive. Then the expected information loss is $\Sigma_i W_i(-\Sigma_j p_{ij} log_2 p_{ij} + \Sigma_j p_{ij} log_2 r_j)$. This value is a function of R, and it is easily proved that this function takes its minimum absolute value when R is the weighted mean where the weights are the probabilities W_i.

This is only a sample case in which the implication of the method is quite clear. In a more general case the weights need not necessarily be comprehended as probabilities. Perhaps the statement that the weighted mean solution minimizes the weighted mean of information loss has value in its own right. This solution will be called the minimum information-loss estimate of probability distribution.

There is another kind of evidence justifying this method. In the above discussion it has been tacitly assumed that none of the information contents degenerates, or, if they do, that they all degenerate in the same subsets. Of course this cannot be assumed generally. If some of the information contents degenerate differently, it is necessary for linear combination that the probability assigned to a subset be divided into each of its constituent states. It is proved that the division proportional to the distribution of r_j in the subset gives the minimum information-loss estimate by solving the resultant system of equations. This is a completely rational solution, because, in the case of an entirely degenerated information content assigning probability one to the whole set, and therefore meaning nothing, the only possible division which does not change the solution by adding this nominal content is the proportional division.

It should be noted, however, that this method is applicable only to a special kind of information-content ensemble—of which all the contents are possibly more or less false. Another extreme case will be found in the ensemble of which contents are more or less incomplete but not false. For this ensemble the solution is obtained by summing $-log\ p_{ij}$ with respect to i for each j (a kind of additivity of amount of information holds in this case!); the resultant probability is obtained from $\Pi_i p_{ij}$ by normalization. As in the previous case, we can define information-loss function as $-\Sigma_j r_j log_2 r_j + \Sigma_j r_j log_2(\Pi_i p_{ij}/\Sigma_j \Pi_i p_{ij})$, and it is easily seen that $r_j = \Pi_i p_{ij}/\Sigma_j \Pi_i p_{ij}$ gives this nonpositive function its maximum value zero.

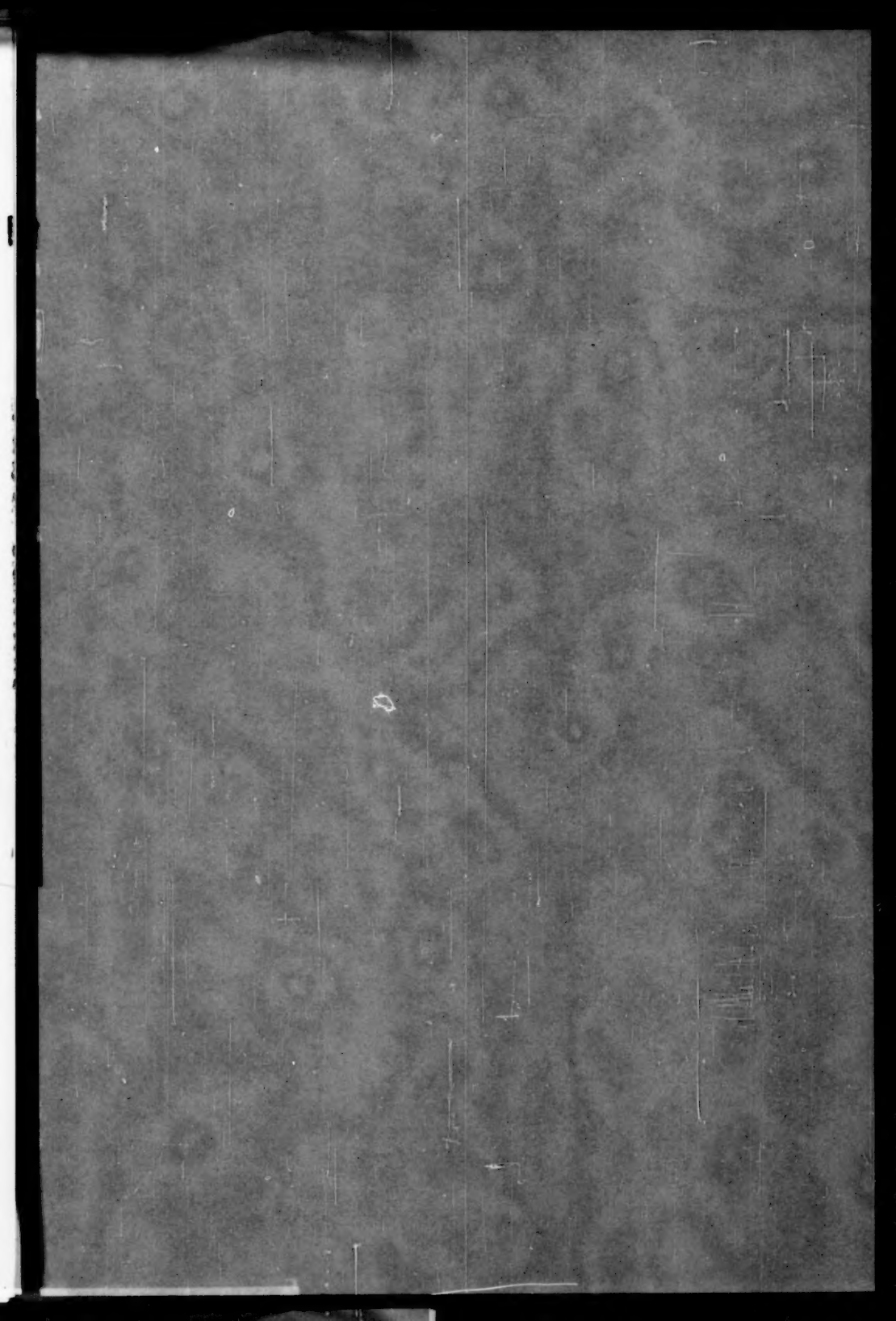
An experiment conducted by the author to detect what rule of information synthesis is adopted by undergraduate students in the latter case revealed a

rather surprising predictability by the above formula for group means. Undoubtedly, this problem of information synthesis deserves the attention not only of experimental psychologists and information theorists but also of many other people who are engaging in the synthesis of information, since obviously they have no rule but intuition for making their syntheses, except when some appropriate statistical techniques are available.

SUMMARY

Communication theory as developed mainly for engineering purposes is incomplete as a theory of actual human communication. The required theory should pay proper attention to the following facts: (*a*) Human communication is a kind of game, sometimes cooperative and sometimes competitive. (*b*) The receiver's rules of decoding are usually not given; they should be developed by the receiver himself for each different information source. (*c*) The receiver may extract more than one information content from a signal. (*d*) The receiver is usually not given objective probabilities, but he estimates them. Accordingly, the amount of information and information content should be defined with respect to subjective probabilities (5). Particular attention should be paid to the problem of information synthesis. Brief consideration has been given to all these topics.

REFERENCES

1. MARSCHAK, J. Toward an economic theory of organization and information. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes*, New York: Wiley, 1954.
2. NEWCOMB, T. M. *Social psychology*. New York: Dryden, 1950.
3. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 197–237.
4. SHANNON, C. E., & WEAVER, W. *The mathematical theory of communication*. Urbana: Univer. of Illinois Press, 1949.
5. TODA, M. Measurement of intuitive probability by a method of game. (Japanese text with English abstract.) *Japanese J. Psychol.*, 1951, **22**, 29–40.
6. TODA, M. Jushinsha to shiteno hito ga jôhô kara handan o kôseisuru mekanizumu ni tsuite. (Man as a receiver, and his mechanism of transforming information to judgment.) In S. Takagi (Ed.), *Shinrigaku ni okeru sûryôka no kenkyû*, Tokyo: Tokyo Univer. Press, 1955. Pp. 131–142.
7. VON NEUMANN, J., & MORGENSTERN, O. *Theory of games and economic behavior*. Princeton: Princeton Univer. Press, 1947.
8. WALD, A. *Statistical decision functions*. New York: Wiley, 1950.

(Received July 19, 1955)